PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ :		(11) International Publication Number	r: WO 98/39446
C12N 15/12, 5/10, 1/21, C07K 14/47, 16/18, C12Q 1/68, G01N 33/50, 33/53, 33/68, A61K 38/17	A2	(43) International Publication Date:	11 September 1998 (11.09.98)

(21) International Application Number:	PCT/US98/04482	
(22) International Filing Date:	6 March 1998 (06.03.98)	
(30) Priority Data:		

7 March 1997 (07.03.97) 60/040,162 US 60/040,333 7 March 1997 (07.03.97) US 7 March 1997 (07.03.97) 60/038,621 US 7 March 1997 (07.03.97) 60/040,161 US 7 March 1997 (07.03.97) 60/040,626 US 60/040.334 7 March 1997 (07.03.97) US 60/040,336 7 March 1997 (07.03.97) US 60/040,163 7 March 1997 (07.03.97) US 11 April 1997 (11.04.97) 60/043,580 US 60/043.568 11 April 1997 (11.04.97) US

(Continued on the following page)

(71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). ROSEN, Craig, A. [US/US], 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). FISCHER, Carrie, L. [US/US]; 5810 Hall Street, Burke, VA 22015 (US). SOP-PET, Daniel, R. [US/US]; 15050 Stillfield Place, Centreville, VA 22020 (US). CARTER, Kenneth, C. [US/US]; 11601 Brandy Hall Lane, North Potomac, MD 20878 (US). BED-NARIK, Daniel, P. [US/US]; 8822 Blue Sea Drive, Columbia, MD 21046 (US). ENDRESS, Gregory, A. [US/US]; 9729 Clagett Farm Drive, Potomac, MD 20854 (US). YU, Guo-Liang [CN/US]; 13524 Straw Bale Lane, Damestown, MD 20878 (US). NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). FENG, Ping [CN/US]; 4 Relda Court, Gaithersburg, MD 20878 (US). YOUNG, Paul, E. [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). GREENE, John, M. [US/US]; 872 Diamond Drive, Gaithersburg, MD 20878 (US). FERRIE, Ann., M. [US/US]; 13203 L Astoria Hill Court, Germantown, MD 20874 (US). DUAN, Roxanne [US/US]; 4541 Fairfield Drive, Bethesda, MD 20814 (US). HU, Jing-Shan [CN/US]; 1247 Lakeside Drive #3034, Sunnyvale, CA 94086 (US). GRAVES, Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place #24, Gaithersburg, MD 20878 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne Terrace #316, Gaithersburg, MD 20878 (US). BREWER, Laurie, A. [US/US]; 14920 Mt. Nebo Road, Poolesville, MD 20837 (US). MOORE, Paul, A. [GB/US]; Apartment #104, 1908 Holly Ridge Drive, McLean, VA 22102 (US). SHI, Yanggu [CN/US]; 437 West Side Drive, Gaithersburg, MD 20878 (US). LAFLEUR, David, W. [US/US]; 1615 Q Street, N.W. #807, Washington, DC 20009 (US). LI, Yi [CN/US]; 1247 Lakeside Drive #3034, Sunnvvale, CA 94086 (US). ZENG, Zhizhen [CN/US]; 13950 Saddleview Drive, Gaithersburg, MD 20878 (US). KYAW, Hla [BU/US]; 520 Sugarbush Circle, Frederick, MD 21703 (US).

- (74) Agents: BROOKES, Anders, A. et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 10850 (US).
- (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM). European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

With an indication in relation to a deposited microorganism furnished under Rule 13bis separately from the description. Date of receipt by the International Bureau: 06 April 1998 (06.04.1998)

(54) Title: 70 HUMAN SECRETED PROTEINS

(57) Abstract

61 34.41 6

The present invention relates to 70 novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.

60/043,314 11 April 1997 (11.04 97 60/043,569 11 April 1997 (11.04 97 60/043,311 11 April 1997 (11.04 97 60/043,671 11 April 1997 (11.04 97 60/043,674 11 April 1997 (11.04 97 60/043,669 11 April 1997 (11.04 97 60/043,312 11 April 1997 (11.04 97 60/043,312 11 April 1997 (11.04 97	US US US US US	60/047.598 23 May 1997 (23.05.97) 60/047,613 23 May 1997 (23.05.97) 60/047.582 23 May 1997 (23.05.97) 60/047.596 23 May 1997 (23.05.97) 60/047,612 23 May 1997 (23.05.97) 60/047,632 23 May 1997 (23.05.97) 60/047,601 23 May 1997 (23.05.97)	US US US US US US US	60/056.882 22 August 1997 (22.08.97) 60/056.637 22 August 1997 (22.08.97) 60/056,903 22 August 1997 (22.08.97) 60/056.888 22 August 1997 (22.08.97) 60/056.880 22 August 1997 (22.08.97) 60/056.890 22 August 1997 (22.08.97)	US US US US US US
60/043,578 11 April 1997 (11.04.97 60/043,570 11 April 1997 (11.04.97 60/043,670 11 April 1997 (11.04.97 60/047,600 23 May 1997 (23.05.97) 60/047,515 23 May 1997 (23.05.97) 60/047,502 23 May 1997 (23.05.97) 60/047,632 23 May 1997 (23.05.97) 60/047,633 23 May 1997 (23.05.97) 60/047,618 23 May 1997 (23.05.97) 60/047,618 23 May 1997 (23.05.97) 60/047,503 23 May 1997 (23.05.97) 60/047,504 23 May 1997 (23.05.97) 60/047,505 23 May 1997 (23.05.97) 60/047,507 23 May 1997 (23.05.97) 60/047,581 23 May 1997 (23.05.97) 60/047,587 23 May 1997 (23.05.97) 60/047,587 23 May 1997 (23.05.97) 60/047,587 23 May 1997 (23.05.97) 60/047,587 23 May 1997 (23.05.97) 60/047,589 23 May 1997 (23.05.97)	US	60/047.585 23 May 1997 (23 05 97) 60/047,586 23 May 1997 (23 05 97) 60/047,590 23 May 1997 (23 05 97) 60/047,591 23 May 1997 (23 05 97) 60/047,592 23 May 1997 (23 05 97) 60/047,593 23 May 1997 (23 05 97) 60/047,591 23 May 1997 (23 05 97) 60/048,974 06 June 1997 (06 06 97) 60/048,964 (06 June 1997 (06 06 97) 60/056,886 22 August 1997 (22 08 97) 60/056,872 22 August 1997 (22 08 97) 60/056,630 22 August 1997 (22 08 97) 60/056,872 22 August 1997 (22 08 97)	US US	60/056,874 22 August 1997 (22.08.97) 60/056,864 22 August 1997 (22.08.97) 60/056,864 22 August 1997 (22.08.97) 60/056,845 22 August 1997 (22.08.97) 60/056,632 22 August 1997 (22.08.97) 60/056,664 22 August 1997 (22.08.97) 60/056,664 22 August 1997 (22.08.97) 60/056,876 22 August 1997 (22.08.97) 60/056,881 22 August 1997 (22.08.97) 60/056,881 22 August 1997 (22.08.97) 60/056,890 22 August 1997 (22.08.97) 60/056,875 22 August 1997 (22.08.97) 60/056,887 22 August 1997 (22.08.97) 60/056,888 22 August 1997 (22.08.97) 60/056,884 22 August 1997 (22.08.97) 60/056,884 22 August 1997 (22.08.97) 60/056,884 22 August 1997 (22.08.97) 60/057,761 05 September 1997 (05.09.97) 60/057,650 05 September 1997 (05.09.97)	US U

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesothe	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovenia
ΑT	Austria	FR	France	LU	Luxembourg	SN	
ΑU	Australia	GA	Gabon	LV	Latvia	SZ	Senegal
ΑZ	Azerbaijar:	GB	United Kingdon.	MC	Monaco	TD	Swaziland
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Chad
BB	Barbados	GH	Ghana	MG	Madagascar		Togo
BE	Belgium	GN	Gumea	MK	The former Yugoslav	TJ TM	Tajikistan
BF	Burkina Fasc	GR	Greece		Republic of Macedonia	TR	Turkmenistar.
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Turkey
BJ	Benin	IE	Ireland	MN	Mongolia		Trinidad and Tobage
BR	Brazil	IL	Israe)	MR	Mauritania	UA UG	Ukraine
BY	Belarus	15	Iceland	MW	Malawi		Uganda
CA	Canada	Τī	Italy	MX	Mexico	US	United States of America
CF	Central African Republic	JP	Japan.	NE	Niger	UZ	Uzbekistar.
CG	Congo	KE	Kenya	NL	Netherlands	VN	Viet Nam
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	YU	Yugoslavia
CI	Côte d'Ivoin	KP	Democratic People's	NZ	New Zealand	zw	Zimbabwe
CM	Cameroon		Republic of Korea	PL.	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstar	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtensteir	SD	Sudan		
DK	Denmark	LK	Sri Lanke	SE	Sweden		
EE	Estonia	LR	Libena	SG	Singapore		

70 Human Secreted Proteins

Field of the Invention

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and their production.

5

10

15

20

25

30

35

Contract Granita

Background of the Invention

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include the commercially valuable human insulin, interferon. Factor VIII, human growth hormone, tissue plasminogen activator, and crythropoeitin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using secreted proteins or the genes that encode them.

10

15

20

25

Summary of the Invention

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig

35

30

WO 98/39446 PCT/US98/04482

3

analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 12301 Park Lawn Drive, Rockville.

Maryland 20852, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

10

15

20

25

30

35

Comment of the second

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl: 0.2M NaH₂PO₄; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 ug/ml salmon sperm blocking DNA: followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a

WO 98/39446 PCT/US98/04482

complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

5

10

15

20

25

30

35

The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single-and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, tor example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation. covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine,

15

20

30

35

5

formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

25 Polynucleotides and Polypeptides of the Invention

FEATURES OF PROTEIN ENCODED BY GENE NO: 1

The translation product of Gene NO: 1 shares sequence homology with alpha-L-fucosidase which is thought to be important as a lysosomal enzyme that hydrolyzes fucose from fucoglycoconjugates. (See Accession No. gi/178409.) Lysosome fructosidase is involved in certain lysosome storage diseases. (See Biochem. Biophys. Res. Commun., 164(1):439-445 (1989).) Fucosidosis, an autosomal recessive lysosomal storage disorder characterized by progressive neurological deterioration and mental retardation. The disease results from deficient activity of alpha-L-fucosidase, a lysosomal enzyme that hydrolyzes fucose from fucoglycoconjugates. This gene likely encodes a novel fucosidase isoenzyme. Based on homology, it is likely that the translated product of this gene is also involved in lysosome catabolism of molecules and

10

15

that aberrations in the concentration and/or composition of this product may be causative in lysosome storage disorders. Preferred polypeptide fragments comprise the amino acid sequence PGHLLPHKWENC (SEQ ID NO: 257).

Gene NO: 1 is expressed primarily in stromal cells, and to a lesser extent in human fetal kidney and human tonsils.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, fucosidosis and other lysosome storage disorders. Similarly, polypeptides and antibodies directed to the polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues of cells, particularly of the nervous system. expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., stromal cells, kidney, tonsils, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology of Gene NO: 1 to alpha-L-fucosidase indicates that polypeptides and polynucleotides corresponding to Gene NO: 1 are useful for the treatment of fucosidosis and general lysosomal disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 134 as residues: Met-1 to Leu-6, Thr-32 to Glu-39, Lys-80 to Lys-85, and Met-90 to Pro-96.

25

30

35

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 2

The translation product of Gene No. 2 shares sequence homology with stromal cell-derived factor-2 (SDF-2) which is a novel secreted factor. See, for example, Gene. 176(1-2):211-214, (1996, Oct. 17.) The amino acid sequence of SDF-2 shows similarity to yeast dolichyl phosphate-D-mannose:protein mannosyltransferases, Pmtlp [Strahl-Bolsinger et al. Proc. Natl. Acad. Sci. USA 90, 8164-8168 (1993)] and Pmt2p [Lussier et al. J. Biol. Chem. 270, 2770-2775 (1995)], whose activities have not been detected in higher eukaryotes. Based on the sequence similarity, the translation product of this gene is expected to share certain biological activities with SDF-2, Pmt1p and Pmt2p.

Gene NO: 2 is expressed primarily in immune system tissue and cancerous tissues, such as liver hepatoma, human B-cell lymphoma, spleen in a patient suffering

trom chronic lymphocytic leukemia, hemangiopericytoma, pharynx carcinoma, breast cancer, thyroid, bone marrow, osteoblasts and to a lesser extent in a few other tissues such as kidney pyramids.

5

10

15

20

25

30

35

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the diseases and conditions which include, but are not limited to, disorders in kidney, liver, and immune organs, particularly cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the kidney, liver, thyroid, and bone marrow expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., liver, spleen, B-cells, pharynx, thyroid, mammary tissue, bone marrow, osteoblasts and kidneys, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology of Gene NO: 2 to stromal cell-derived factor-2 indicates that polypeptides and polynucleotides corresponding to Gene NO: 2 are useful for diagnosis and therapeutic treatment of disorders in kidney, liver, and immune organs since stromal cells play important role in organ function. Stroma carries the blood supply and provides support for the growth of parenchymal cells and is therefore crucial to the growth of a neoplasm. Nucleic acids of the present invention comprise, but preferably do not consist of, and more preferably do not comprise, SEQ ID NO: 3 from US Patent No. 5,576.423, incorporated herein by reference, and shown herein as SEQ ID NO: 258).

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 135 as residues: His-56 to Gly-65, Ala-74 to Ser-80, Ile-84 to Pro-97, Leu-124 to Glu-129, Glu-135 to Asp-143, Gly-175 to Ser-180, and Ala-194 to Thr-199.

FEATURES OF PROTEIN ENCODED BY GENE NO: 3

The translation product of Gene NO: 3 shares sequence homology with LZIP-1. LZIP-2 and other leucine zipper proteins, which are thought to be important in nucleic acid binding. This gene has been reported in Mol. Cell. Biol. 17 (9), 5117-5126 (1997) as "Luman". Luman is a cyclic AMP response element (CRE)-binding protein/activating transcription factor 1 protein of the basic leucine zipper superfamily. It binds CREs in

1.5

20

25

30

35

Exercise to the second

۶

vitro and activates CRE-containing promoters when transfected into COS7 cells. The complete amino acid sequence of Luman reported in Mol. Cell. Biol. 17 (9): 5117-5126 (1997) is:

MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLDWALPLSEVPSDWEVDDLL CSLLSPPASLNILSSSNPCLVHHDHTYSLPRETVSMDLESESCRKEGTQMTPQH MEELAEQEIARLVLTDEEKSLLEKEGLILPETLPLTKTEEQILKRVRRKIRNKRSA QESRRKKKVYVGGLESRVLKYTAQNMELQNKVQLLEEQNLSLLDQLRKLQAM VIEISNKTSSSSTCILVLLVSFCLLLVPAMYSSDTRGSLPAEHGVLSRQLRALPSE DPYQLELPALQSEVPKDSTHQWLDGSDCVLQAPGNTSCLLHYMPQAPSAEPPL EWPFPDLSS EPLCRGPILPLQANLTRKGGWLPTGSPSVILQDRYSG (SEQ ID N:259).

Gene NO: 3 is expressed primarily in apoptotic T-cells and Soares senescent cells and to a lesser extent in multiple tissues and cell types, including, multiple sclerosis tissue, and hippocampus.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immunologically mediated disorders, transplantation, immunodeficiency, and tumor necrosis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and transplantation, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., multiple sclerosis tissue, hippocampus, bone marrow and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology of Gene NO: 3 to leucine zipper nucleic acid binding proteins indicates that polypeptides and polynucleotides corresponding to Gene NO: 3 are useful for diagnosis and treatment of immunologically mediated disorders, transplantation, immunodeficiency, and tumor necrosis. The secreted nucleic acid binding protein in the apoptotic tissues may be involved in the disposal of the DNA released by apoptotic cells. Furthermore, the studies conducted in support of Luman suggest that the translation product of this gene may be used to identify transcriptional regulation elements which in turn are useful in modulation of immune function.

15

20

25

30

35

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 136 as residues: Asn-7 to Ser-12, Tyr-32 to Gly-38, Pro-55 to Tyr-60, Glu-70 to Thr-76, and Pro-104 to Leu-110.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 4

The translation product of Gene NO: 4 shares sequence homology with a number of tetraspan transmembrane surface molecules such as human metastasis tumor suppressor gene, CO-029 tumor associated antigen protein, CD53 hematopoietic antigen, human membrane antigen TM4 superfamily protein, metastasis controlling peptide, and human CD9 sequence, which are thought to be important in development of cancer, immune system development and functions.

Gnee NO: 4 is expressed primarily in cancers of several different tissues and to a lesser extent in normal tissue like prostate, skin and kidney.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancers and disorders of the immune system, prostate and kidney. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the kidney, skin, prostate and immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., kidney, skin and prostate, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology of Gene NO: 4 to tetraspan transmembrane surface molecules such as human metastasis tumor suppressor gene. CO-029 tumor associated antigen protein. CD53 hematopoietic antigen, human membrane antigen TM4 superfamily protein, metastasis controlling peptide, and human CD9 sequence, indicates that polypeptides and polynucleotides corresponding to Gene NO: 4 are involved with the cellular control of growth and differentiation. Therefore, the translation product of this gene is believed to be useful for diagnosis and treatment of neoplasia and disorders of the kidney, skin and prostate. For example, recombinant protein can be produced in transformed host cells for diagnostic and prognostic applications. Alterations in the protein sequence are indicative of the presence of

10

15

20

25

30

35

malignant cancer, or of a predisposition to malignancy, in a subject. Gene therapy can be used to restore the wild-type gene product to a subject. Additionally, the antibodies are a useful tool for the identification of hematopoietic neoplasms, and may prove helpful for identifying morphologically poorly defined cells. Moreover, this protein can be used to isolate cognate receptors and ligands and identify potential agonists and antagonists using techniques known in the art. The protein also has immunomodulatory activity, regulates hematopoiesis and stimulates growth and regeneration as a male/female contraceptive, increases fertility depending on activin and inhibin like activities. Other uses are as a chemotactic agent for lymphocytes, treatment of coagulation disorders, an anti-inflammatory agent, an antimicrobial or analgesic and as a modulator of behavior and metabolism. The DNA can be used in genetic diagnosis or gene therapy, and for the production of recombinant protein. It can also be used to identify protein expressing cells, isolate related sequences, prepare primers for genetic fingerprinting and generate anti-protein or anti-DNA antibodies. In addition, residues 1-71, in the translation product for this gene are believed to be the extracellular domain. Thus, polypeptide comprising residues 1-71 or derivatives (including fragments) or analogs thereof, are useful as a soluble polypeptide which may be routinely used therapeutically to antagonize the activities of the receptor.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 137 as residues: Lvs-118 to Phe-127, Asn-145 to Ala-160, and Thr-177 to Val-188.

FEATURES OF PROTEIN ENCODED BY GENE NO: 5

Gene NO: 5 is expressed primarily in human testes.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the testes including cancer and reproductive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., testes and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of Gene NO: 5 indicates that the protein product of this gene is useful for treatment/diagnosis of diseases of the testes, particularly testicular cancer since expression is observed primarily in the testes.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 138 as residue: Gly-22 to Gln-30.

FEATURES OF PROTEIN ENCODED BY GENE NO: 6

5

10

15

20

25

30

35

The translation product of Gene NO: 6 shares sequence homology with GALNS (N-acetylgalactosamine 6-sulphatase) which is thought to be important in the storage of the glycosaminoglycans, keratan sulfate and chondroitin 6-sulfate. See Genbank accession no. gil618426. Based on the sequence similarity, the translation product of this gene is expected to share biological activities with GALNS.

Gene NO: 6 is expressed primarily in human bone marrow.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, storage disorders of glycosaminoglycans, keratan sulfate and chondroitin 6-sulfate, e.g., Morquio A syndrome. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly involving cell storage disorder, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., bone marrow and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology of Gene NO: 6 to N-acetylgalactosamine 6-sulphatase indicates that polypeptides and polynucleotides corresponding to Gene NO: 6 are useful for the treatment and diagnosis of storage disorders of glycosaminoglycans, keratan sulfate and chondroitin 6-sulfate. Such disorders are known in the art and include, e.g., Morquio A syndrome which is caused by an error of mucopolysaccharide metabolism with excretion of keratan sulfate in urine. Morquio A syndrome is characterized by severe skeletal defects with short stature, severe deformity of spine and thorax, long bones with irregular epiphyses but with shafts of normal length, enlarged joints, flaccid ligaments, and waddling gait; autosomal recessive inheritance.

10

15

20

25

30

35

GERGAGE AS

Professional Contraction

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 139 as residues: Gly-29 to Pro-36 and Glu-57 to Leu-64.

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

The translation product of Gene NO: 7 shares sequence homology with carboxy peptidase E and H (carboxypeptidase E is thought to be important in the biosynthesis of numerous peptide hormones and neurotransmitters). The translation product of this gene also shares sequence homology with bone-related carboxypeptidase "OSF-5" from the mouse. See European patent application EP-588118-A. Based on the sequence similarity to OSF-5, the translation product of this gene will hereinafter sometimes be referred to as "human-OSF-5" or "hOSF-5".

Gene NO: 7 is expressed primarily in tumor cell lines derived from connective tissues including chondrosarcoma, synovial sarcoma. Wilm's tumor and rhabdomyosarcoma and to a lesser extent in a myeloid progenitor cell line, bone marrow, and placenta.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, various cancers involving the skeletal system and connective tissues in general, in particular at cartilage interfaces. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system and various other tumor tissues, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., connective tissues and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The restricted tissue distribution and homology of Gene NO: 7 to carboxypeptidase E and mouse OSF-5 indicates that polypeptides and polynucleotides corresponding to Gene NO: 7 are for processing of peptides to their mature form that may have various activities similar to the activities of neuropeptides but in the periphery. In addition the abundance of expression in cancer tissues indicates that aberrant expression and subsequent processing may play a role in the progression of malignancies, e.g., growth factor and/or adhesion factor activities. In particular, the expression of this gene is restricted to connective tissues and embryonic tissues.

15

20

25

30

35

Furthermore, it is overexpressed in cancers of these same tissues (i.e., in sarcomas). Moreover, hOSF-5 shares very strong sequence similarity with mOSF-5 which is a known bone growth factor and is thought to be useful in obtaining products for the diagnosis and treatment of bone metabolic diseases, e.g., osteoporosis and Paget's disease. Like OSF-5, the translation product of this gene is believed to be a bone-specific carboxypeptidase which acts as an adhesion molecule/growth factor and takes part in osteogenesis at the site of bone induction. hOSF-5 can, therefore, be used to treat bone metabolic diseases, osteoporosis, Paget's disease, osteomalacia, hyperostosis or osteopetrosis. Furthermore, hOSF-5 can be used to stimulate the regeneration of bone at the site of mechanical damage, e.g., accidentally or surgically caused fractures.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 140 as residues: Leu-24 to Val-30. Ala-89 to Lys-94. Phe-150 to Trp-157, Leu-162 to Asp-167, Asp-187 to Ser-199. His-241 to Asp-254, and Pro-362 to Asp-376.

FEATURES OF PROTEIN ENCODED BY GENE NO: 8

Gene NO: 8 is expressed primarily in bone marrow, and to a lesser extent in an erythroleukemia cell line.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematological disorders including cancer and anemia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematologic systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., bone marrow, kidney, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: δ are useful as a growth factor for hematopoietic stem cells or progenitor cells, e.g., in the treatment of bone marrow stem cell loss in chemotherapy patients and in the treatment of kidney disease.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 141 as residues: Gly-30 to Lys-35.

15

20

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

Gene NO: 9 is expressed primarily in neutrophils.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the cell type present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to. inflammatory diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the cell type indicated. For a number of disorders of the above tissues or cells. particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues or cell types (e.g., neutrophils, bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.c., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 9 are useful for immune modulation or as a growth factor to stimulate neutrophil differentiation or proliferation that may be useful in the treatment of neutropenia.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 142 as residues: Thr-22 to Pro-37.

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

Gene NO: 10 is expressed primarily in the epidermis.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the epidermis such as psoriasis or eczema or may be involved in the normal proliferation or differentiation of the epithelial cells or fibroblasts constituting the skin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skin, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., epidermis and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder.

relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 10 are useful for diagnosis and treatment of skin conditions and as an aid in the healing of various epidermal injuries including wounds, and diabetic ulcers.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 143 as residues: Ser-3 to Ser-9 and Trp-27 to Glu-32.

FEATURES OF PROTEIN ENCODED BY GENE NO: 11

10

15

20

25

30

35

The translation product of Gene NO: 11 shares sequence homology with phosphatidylcholine 2-acylhydrolase (PLA2). See, for example, Genbank accession no. gil190004. PLA2 is involved in inflammation, where it is responsible for the conversion of cell membrane phospholipids into arachidonic acid. Arachidonic acid in turn feeds into both the lipoxygenase and cyclooxygenase pathways to produce leukotrienes (involved in chemotaxis, vasoconstriction, bronchoconstriction, and increased vascular permeability) and prostaglandins (responsible for vasodilation, potentiate edema, and increased pain). Diseases in which PLA2 is implicated as a major factor include rheumatoid arthritis, sepsis, ischemia, and thrombosis. The inventors refer to the translation product of this gene as PLA2-like protein based on the sequence similarity. Furthermore, owing to the sequence similarity PLA2 and PLA2-like protein are expected to share certain biological activities.

Gene NO: 11 is expressed primarily in human cerebellum and in T-cells.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cerebellum disorders, rheumatoid arthritis, sepsis, ischemia, and thrombosis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cerebellum and Purkinje cells, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., brain, bone marrow, T-cells, and cancerous and wounded tissues) or bodily fluids (e.g., scrum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 11 are useful for diagnosis and treatment of cerebellum disorders, rheumatoid arthritis, sepsis, ischemia, and thrombosis. This gene is also useful as a chromosome marker. It is believed to map to Chr.15, D15S118-D15S123.

5

10

15

20

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 12

Gene NO: 12 is expressed primarily in highly vascularized tissues such as placenta, uterus, tumors, fetal liver, fetal spleen and also in the C7MCF7 cell line treated with estrogen.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endometriosis, endometritis, endometrial carcinoma, primary hepatocellular carcinoma, and spleen-related diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endometrium, liver and spleen, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., endometrium, liver, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 12 are useful for diagnosis and treatment of diseases of the endometrium (such as endometrial carcinoma, endometriosis, and endometritis), liver diseases (such as primary hepatocellular carcinoma), and spleen-related diseases.

SEQ ID NO: 145 as residues: Ala-29 to Leu-35, Leu-50 to Ser-57. Glu-96 to Glu-105, Asp-140 to Asp-148, and Asn-191 to Ser-197.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 13

Gene NO: 13 is expressed primarily in B cell lymphoma and to a lesser extent in other tissues.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, B cell lymphoma; hematopoietic disorders; immune dysfunction.

WO 98/39446

5

10

15

20

30

35

PCT/US98/04482

17

Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., bone marrow and B-cells and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Enhanced expression of this gene product in B cell lymphoma indicates that it may play a role in the proliferation of hematopoietic cells. It is also believed to be involved in the survival and/or differentiation of various hematopoietic lineages. Expression in lymphoma also indicates that it may be involved in other cancers and abnormal cellular proliferation. The tissue distribution, therefore, indicates that polypeptides and polynucleotides corresponding to Gene NO: 13 are useful for the diagnosis and/or therapeutic treatment of hematopoietic disorders, particularly B cell lymphoma. Furthermore, since overexpression of this gene is associated with the development of B cell lymphoma, antagonists of this protein are useful to interfere with the progression of the disease. This protein is useful in assays for identifying such antagonists. Assays for identifying antagonists are known in the art and are described briefly elsewhere herein. Preferred antagonists include antibodies and antisense nucleic acid molecules. Preferred are antagonists which inhibit B-cell proliferation.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 14

The translation product of Gene NO: 14 shares sequence homology with very low density lipoprotein receptor which is thought to be important in transport of lipoproteins. Owing to the sequence similarity the translation product of this gene is believed to share certain biological activities with VLDL receptors. Assaying such activity may be achieved by assays known in the art and set forth elsewhere herein.

This gene is expressed primarily in human synovium, umbilical vein endothelial cells, CD34+ cells, Jurkat cells, and HL60 cells, and to a lesser extent in thymus, meningioma, hypothalmus, adult testis, and fetal liver and spleen.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, atherosclerosis, ataxia malabsortion, vascular damage, hyperlipidemia, and other cardiovascular diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular and hematological systems. expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., endothelium, thymus meningioma, hypothalmus, testes, liver, and spleen and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in the vascular endothelial cells and homology to VLDL receptors indicates that polypeptides and polynucleotides corresponding to Gene NO: 14 are useful for diagnosis and treatment of atherosclerosis, ataxia malabsortion, and hyperlipidemia. These and other factors often result in other cardiovascular diseases. Additionally, the presence of the gene product in cells of blood lineages indicates that it may be useful in hematopoietic regulation and hemostasis.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 147 as residues: Pro-39 to Ser-52, Trp-71 to Thr-76, and Pro-94 to His-100.

20

5

10

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 15

The translation product of Gene NO: 15 shares sequence homology with kallikrein which is thought to be important in blood pressure and renal secretion. Furthermore, this gene has now been characterized as a novel hepatitis B virus X binding protein that inhibits viral replication. See, for example, J. Virol. 72 (3), 1737-1743 (1998).

This gene is expressed primarily in kidney, placenta, lung, aorta and other endothelial cells, caudate nucleus and to a lesser extent in melanocytes, liver, adipose tissue.

30

35

25

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, renovascular hypertension, renal secretion, electrolyte metabolism, toxemia of pregnancy. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renovascular or respiratory vascular systems, expression of this gene

10

15

20

25

30

35

The Control of the Control of the

at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., kidney, placenta, lung, endothelial cells, melanocytes, liver, and adipose tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to kallikrein indicates that polypeptides and polynucleotides corresponding to Gene NO: 15 are useful for treating renovascular hypertension, renal secretion, electrolyte metabolism, toxemia of pregnancy and hydronephrosis. The protein expression in the organs like kidney, lung and vascular endothelial cells indicates the gene involvement in hemodynamic regulatory functions. The translation product of this gene is also useful in the treatment of viral infection, particularly liver infection, and particularly hepatitis B virus(es).

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 148 as residues: Leu-9 to Asn-15 and Thr-56 to Asp-61.

FEATURES OF PROTEIN ENCODED BY GENE NO: 16

The translation product of Gene NO: 16 shares sequence homology with secretory component protein, immunoglobulins and their receptors which are thought to be important in immunological functions. The amino acid sequence of secretory component protein can be accessed as accession no. pirlA02112, incorporated herein by reference.

Gene NO: 16 is expressed primarily in macrophages, monocytes and dendritic cells and to a lesser extent in placenta and brain.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation and tumors. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues or cells (e.g., macrophages, monocytes, dendritic cells, plancenta and brain, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a

disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to immunoglobulins and secretory component protein indicates that polypeptides and polynucleotides corresponding to Gene NO: 16 are useful for diagnosis and treatment of inflammation and bacterial infection, and other diseases where immunomodulation would be beneficial.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 149 as residues: Pro-37 to Cys-51, Gln-53 to Cys-60, Asn-99 to Gly-106, Gly-145 to Glu-151, and Ile-159 to Ser-164.

10

15

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 17

The translation product of Gene NO: 17 is evolutionarily conserved and shares sequence homology with proteins from yeast and *C. elegans*. See, for example. Genbank accession no.gil746540. As is known in the art, strong sequence similarity to a secreted protein from C. elegans is predictive of cellular location of human proteins.

Gene NO: 17 is expressed primarily in colon carcinoma cell lines, messangial cells, many tumors like T cell lymphoma, osteoclastoma. Wilm's tumor, adrenal gland tumor, testes tumor, synovial sarcoma, and to a lesser extent in placenta, lung and brain.

20

25

30

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, rapidly growing/dividing cells such as cancerous tissue, including, colon carcinoma, lymphomas, and sarcomas. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the gastrointestinal, hematological and immune systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., placenta, lung, brain, colon, messangial cells, adrenal gland, T-cells, testes, and lymph tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

35

EASTE TOWN

The tissue distribution in colon cancer and many other tumors indicates that the polynucleotides and polypeptides of Gene NO: 17 are useful for cancer diagnosis and therapeutic targeting. The extracellular nature may contribute to solid tumor

10

15

20

25

30

35

immunosuppression, angiogenesis and cell growth stimulation. The tissue distribution of this gene in cells of the immune system indicates that polypeptides and polynucleotides corresponding to Gene NO: 17 are useful for treatment, prophylaxis and diagnosis of immune and autoimmune diseases, such as lupus, transplant rejection, allergic reactions, arthritis, asthma, immunodeficiency diseases, leukemia, and AIDS. Its expression predominantly in hematopoietic cells also indicates that the gene could be important for the treatment and/or detection of hematopoietic disorders such as graft versus host reaction, graft versus host disease, transplant rejection, myelogenous leukemia, bone marrow fibrosis, and myeloproliferative disease. The protein can also be used to enhance or protect proliferation, differentiation and functional activation of hematopoietic progenitor cells such as bone marrow cells, which could be useful for cancer patients undergoing chemotherapy or patients undergoing bone marrow transplantation. The protein may also be useful to increase the proliferation of peripheral blood leukocytes, which could be useful in the combat of a range of hematopoietic disorders including immunodeficiency diseases, leukemia, and septicemia.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 150 as residues: Val-131 to Asn-136.

FEATURES OF PROTEIN ENCODED BY GENE NO: 18

The translation product of Gene NO: 18 shares sequence homology with immunoglobulin, which is thought to be important in immunoreactions.

Gene NO: 18 is expressed primarily in macrophage.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., macrophage and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.c., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in macrophages and the weak homology to immunoglobin indicates that polypeptides and polynucleotides corresponding to Gene

10

15

20

25

30

35

PROTECTION OF HUMBER AS

NO: 18 are useful for diagnosing and treating immune response disorders, including inflammation, antigen presentation and immunosurveillance.

FEATURES OF PROTEIN ENCODED BY GENE NO: 19

The translation product of Gene NO: 19 shares sequence homology with proline rich proteins which are thought to be important in protein-protein interaction.

This gene has a wide range of tissue distribution, but is expressed primarily in normal prostate, synovial fibroblasts, brain amygdala depression, fetal bone and fetal cochlea, and to a lesser extent in adult retina, umbilical vein endothelial cells, atrophic endometrium, osteoclastoma, melanocytes, pancreatic carcinoma and smooth muscle.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer metastasis, wound healing, tissue repair. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal, connective tissues, reproductive and central nervous system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., brain, prostrate, fibroblasts, bone, cochlea, retina, endothelial cells, endometrium, pancreas and smooth muscle, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to proline-rich proteins indicates that the protein is a extracellular matrix protein or an ingredient of bodily fluid. Polypeptides and polynucleotides corresponding to Gene NO: 19 are useful for cancer metastasis intervention, tissue culture additive, bone modeling, wound healing and tissue repair.

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

Gene NO: 20 is expressed primarily in prostate cancer, leukocytes, meningima, adult liver, pancreas, brain, and to a lesser extent in lung.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, prostate cancers. Similarly, polypeptides and antibodies directed to these

WO 98/39446 PCT/US98/04482

23

polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate and brain, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., prostate, leukocytes, memingima, liver, brain, pancreas and lung, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Prostate cancer cell lines are known to be responsive to estrogen and androgen. The protein expression of Gene NO: 20 appears to be influenced by both estrogen and androgen levels. The prostate cancer tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 20 are is useful in the intervention and detection of prostate hyperplasia and prostate cancer.

15

20

25

30

35

10

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

The translation product of Gene NO: 21 is identical to the human wnt-7a genc. Wnt-7a is a secreted signaling molecule, thought to be important in signaling and the regulation of cell fate and pattern formation during embryogenesis. Specifically, knock out studies in mice have demonstrated that wnt7a plays a critical role in the development of the dorsal-ventral patterning in the developing limb, and to a lesser extent plays a role in the development of anterior-posterior patterning. Overexpression of wnt7a can induce transformation of cultured mammary cells, suggesting that it is an oncogene.

Expression of Gene NO: 21 has only been observed in testes.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, testicular cancer; abnormal limb development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the testes or developing embryo. For a number of disorders of the above tissues or cells, particularly of the developing embryo, expression of this gene at significantly higher or lower levels may routinely be detected in the developing embryo or amniotic fluid taken from a pregnant individual and compared relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Also, expression of this gene at significantly higher or lower levels may routinely be detected in the testes of patient suffering from testicular cancer and compared relative to the

standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to mouse wnt7a indicates that polypeptides and polynucleotides corresponding to Gene NO: 21 are useful to restore abnormal limb development in an affected individual. Furthermore, its oncogenic potential and tissue distribution indicates that it could serve as a diagnostic for testicular cancer.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 154 as residues: Gly-22 to Arg-28.

10

15

20

25

30

35

PARTICL ST. SWC 989-PARE

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 22

Gene NO: 22 is expressed primarily in fetal liver/spleen, breast, testes and placenta and to a lesser extent in brain, and a series of cancer tissues.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, brain diseases, male infertility, and disposition to pregnant miscarriages. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, hematopoietic system, and sexual organs, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., liver, spleen, testes, placenta, and brain, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene indicates that polypeptides and polynucleotides corresponding to Gene NO: 22 are useful as a marker for non-differentiated, dividing cells and hence could serve as an oncogenic marker. Its high expression in fetal liver, suggests an involvement in hematopoiesis and/or the immune system. Hence it is useful as a factor to enhance an individuals immune system, e.g., in individuals with immune disorders. It is also thought to affect the survival, proliferation, and differentiation of a number of hematopoietic cell lineages, including hematopoietic stem cells. Its disruption, e.g., mutation or altered expression, may also be a marker of immune disorder. Its expression in the testes, suggests it may be important in controlling male fertility. Expression of this gene in breast further reflects a

role in immune function and immune surveillance (breast lymph node). This gene is believed to be useful as a marker for breast cancer.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 155 as residues: Gln-57 to Lvs-70 and Ala-91 to Pro-100.

5

10

15

20

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 23

Gene NO: 23 is expressed primarily in bone marrow and brain (whole and fetal).

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, immune and hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous and hematopoietic systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., bone marrow, brain, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 23 are useful in the diagnosis and treatment of disorders related to the central nervous system (e.g. neuro-degenerative conditions, trauma, and behavior abnormalities) and hematopoiesis. In addition, the expression in fetal brain indicates a role for this gene product in diagnosis of predisposition to developmental defects of the brain.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 156 as residues: Thr-23 to Tyr-29.

30

35

THE SECURE OF YOUR AS

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

Gene NO: 24 is expressed primarily in smooth muscle, placenta, prostate, and osteoblasts.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cardiovascular pathologies. Similarly, polypeptides and antibodies

15

20

25

30

35

directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular, reproductive and skeletal systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., placenta, smooth muscle, prostrate, and osteoblasts, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 24 are useful for detection and treatment of neoplasias and developmental abnormalities associated with these tissues.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 157 as residues: Asn-21 to Thr-26.

FEATURES OF PROTEIN ENCODED BY GENE NO: 25

The translation product of Gene NO: 25 shares sequence homology with Pregnancy Associated Mouse Protein (PAMP)-1. (See, FEBS Lett 1993 May 17;322(3):219-222). Based on the sequence similarity the translation product of this gene is expected to share certain biological activities with PAMP-1.

Gene NO: 25 is expressed primarily in 12-week-old human embryos and prostate.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, prostate disorders (cancer). Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., embryonic tissue, and prostate, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10

15

20

25

30

35

that of two governors

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 25 are useful for the diagnosis and treatment of prostate disorders (such as cancer) and developmental abnormalities and fetal deficiencies. The homology to PAMP-1 indicates that this gene and gene product are useful in detecting pregnancy.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 158 as residues: Pro-23 to Glu-28 and Ser-44 to Gly-55.

FEATURES OF PROTEIN ENCODED BY GENE NO: 26

Gene NO: 26 is expressed primarily in testes and to a lesser extent in epididymis.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive and endocrine disorders, as well as testicular cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive and endocrine systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., testes, and epididymis, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 26 are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g., endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to by useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product may be expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications.

10

15

20

25

30

35

E1.50000 (A. C. 954-444-4)

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 159 as residues: Pro-24 to Gly-33 and Arg-70 to Gly-76.

FEATURES OF PROTEIN ENCODED BY GENE NO: 27

The translation product of Gene NO: 27 shares sequence homology with salivary protein precursors which are thought to be important in immune response and production of secreted proteins.

Gene NO: 27 is expressed primarily in salivary gland tissue.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, diseases of the salivary gland. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, digestive system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., salivary gland, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to salivary secreted protein indicates that polypeptides and polynucleotides corresponding to Gene NO: 27 are useful for treatment of immune disorders and diagnostic uses related to secretion of protein in disease states. For example, the gene product can be used as an anti-microbial agent, an ingredient for oral or dental hygiene, treatment of xerostomia, sialorrhea, intervention for inflammation including parotitis, and an indication for tumors in the salivary gland (adenomas, carcinomas).

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 160 as residues: Asp-21 to Gly-28, Asp-30 to Glu-43. Glu-49 to Glu-62, and Thr-75 to Pro-83.

FEATURES OF PROTEIN ENCODED BY GENE NO: 28

Gene NO: 28 is expressed primarily in human fetal heart tissue and to a lesser extent in olfactory tissue.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, olfactory and cardiovascular disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, olfactory and vascular systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., olfactory tissue, and heart, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 28 are useful for diagnosis and treatment of immune, olfactory and vascular disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 161 as residues: Cys-33 to Gly-44, Arg-71 to Arg-78, Ser-130 to Gly-142, Lys-150 to Gly-157, and Thr-159 to Asp-177.

FEATURES OF PROTEIN ENCODED BY GENE NO: 29

5

10

15

20

25

30

35

Gene NO: 29 is expressed primarily in brain and to a lesser degree in activated macrophages, endothelial and smooth muscle cells, and some bone cancers.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of brain and endothelial present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegeneration, inflammation and other immune disorders, fibrotic conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification brain, smooth muscle, and endothelium. For a number of disorders of the above tissues or cells, particularly of the brain and endothelium, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues or cell types (e.g., brain, endothelial cells, macrophages, smooth muscle, and bone, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10

15

20

25

30

35

14.41 1 2.WC 989944A2

Tissue distribution suggests polypeptides and polynucleotides corresponding to Gene NO: 29 are useful in study and treatment of neurodegenerative and immune disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 162 as residues: Asn-18 to Glu-20, Ser-33 to Gln-48, Cys-55 to Ser-56, Pro-67 to Cvs-69.

FEATURES OF PROTEIN ENCODED BY GENE NO: 30

Gene NO: 30 is expressed primarily in early stage human brain and to a lesser extent in cord blood, heart, and some tumors.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of developing CNS tissue present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cardiovascular and neurodegenerative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and immune systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., brain and heart, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that that polypeptides and polynucleotides corresponding to Gene NO: 30 are useful for the treatment of cancer and of neurodegenerative and cognitive disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 31

Gene NO: 31 is expressed primarily in brain and thymus and to a lesser extent in several other organs and tissues including the hematopoietic system, liver skin and bone

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, CNS disorders, hematopoietic system disorders, disorders of the endocrine system, bone, and skin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential

10

20

25

30

35

144520040 NA . .

identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly CNS disorders, hematopoietic system disorders, disorders of the endocrine system, bone, and skin, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g.,

hematopoietic cells, brain, thymus, liver, bone, and epidermis, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 31 are useful for treatment and diagnosis of CNS disorders, hematopoietic system disorders, disorders of the endocrine system, and of bone and skin.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 164 as residues: Thr-35 to Arg-40, Pro-55 to His-75, Pro-93 to Ala-98, Ala-111 to Pro-119, and Pro-132 to Glu-138.

FEATURES OF PROTEIN ENCODED BY GENE NO: 32

Gene NO: 32 is expressed primarily in organs and tissue of the nervous system and to a lesser extent in various developing tissues and organs.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the central nervous system and disorders of developing and growing tissues and organs. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly disorders of the CNS, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., tissue of the nervous system and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 32 are useful for diagnosis and treatment of disorders of the central nervous system, general neurological diseases and neoplasias.

10

15

20

25

30

35

4450000 DWG 100000

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 165 as residues: Ser-33 to Lys-41 and Glu-86 to Glu-91.

FEATURES OF PROTEIN ENCODED BY GENE NO: 33

Residues 141-156 in the translation product for Gene NO: 33 as shown in the sequence listing matches phosphopantetheine binding site motifs. Phosphopantetheine (or pantetheine 4' phosphate) is the prosthetic group of acyl carrier proteins (ACP) in some multienzyme complexes where it serves as a 'swinging arm' for the attachment of activated fatty acid and amino-acid groups. Phosphopantetheine is attached to a serine residue in these proteins. ACP proteins or domains have been found in various enzyme systems which are listed below. Fatty acid synthetase (FAS), which catalyzes the formation of long-chain fatty acids from acetyl-CoA, malonyl-CoA and NADPH. Bacterial and plant chloroplast FAS are composed of eight separate subunits which correspond to the different enzymatic activities; ACP is one of these polypeptides. Fungal FAS consists of two multifunctional proteins, FAS1 and FAS2; the ACP domain is located in the N-terminal section of FAS2. Vertebrate FAS consists of a single multifunctional enzyme; the ACP domain is located between the beta-ketoacyl reductase domain and the C-terminal thioesterase domain. Based on the presence of a phosphopantetheine binding site in the translation product of this gene, it is believed to share activities fatty acid synthetase polypeptides. Such activities may be assayed by methods known in the art.

This gene is expressed primarily in developing and rapidly growing tissues like placenta fetal heart and endometrial tumor and to a lesser extent in B and T cell lymphoma tissues

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and disorders of developing tissues and organs. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic tissues and developing organs and tissues, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., embryonic tissue, endometrium, B-cells, and T-cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to

15

20

25

30

35

Committee Committee

the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 33 are useful for treatment and diagnosis of cancer in the hematopoietic system developing organs and tissues. It may also be useful for induction of cell growth in disorders of the hematopoietic system and other tissue and organs. The homology to fatty acid synthetases indicates that this gene product is useful in the diagnosis and treatment of lipid metabolism disorders such as hyperlipidemia.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 166 as residues: Arg-27 to Glu-34.

FEATURES OF PROTEIN ENCODED BY GENE NO: 34

Gene NO: 34 is expressed primarily in breast and testes tissues and to a lesser extent in hematopoietic tissues including tonsils, T cells and monocytes.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the reproductive organs and systems, including cancer, autoimmune diseases and inflammatory diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive organs and hematopoietic tissues, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., hemotopoietic cells, T-cells and monocytes, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Nucleic acids comprising sequence of this gene are also useful as chromosome markers since this gene maps to Chr.15, D15S118-D15S123.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 34 are useful for treatment of diseases of the reproductive organs and hematopoietic system including cancer, autoimmune diseases and inflammatory diseases.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 167 as residues: Phe-81 to Lys-86.

10

15

20

25

30

35

ENCT 11 14 WELLER

FEATURES OF PROTEIN ENCODED BY GENE NO: 35

The translation product of Gene NO: 35 shares sequence similarity with the mouse cytokine-inducible inhibitor of signaling. See, e.g., Nature 1997 Jun 26:387(6636):917-921. Cytokines are secreted proteins that regulate important cellular responses such as proliferation and differentiation. Key events in cytokine signal transduction are well defined: cytokines induce receptor aggregation, leading to activation of members of the JAK family of cytoplasmic tyrosine kinases. In turn, members of the STAT family of transcription factors are phosphorylated, dimerize and increase the transcription of genes with STAT recognition sites in their promoters. Less is known of how cytokine signal transduction is switched off. Expression of the mouse SOCS-1 protein inhibited both interleukin-6- induced receptor phosphorylation and STAT activation. We have also cloned two relatives of SOCS-1, named SOCS-2 and SOCS-3, which together with the previously described CIS form a new family of proteins. Transcription of all four SOCS genes is increased rapidly in response to interleukin-6, in vitro and in vivo, suggesting they may act in a classic negative feedback loop to regulate cytokine signal transduction. The translation product of this gene is believed to have similar biological activities as this family of mouse genes. The biological activity of the translation product of this gene may be assayed by methods shown in Nature 1997 Jun 26;387(6636): 917-921, which is incorporated herein by reference in its entirety.

Gene NO: 35 is expressed primarily in tissues of hematopoietic origin including activated monocytes, neutrophils, activated T-cells and to a lesser extent in breast, adipose tissue and dendritic cells.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the hematopoietic system including cancer autoimmune diseases and inflammatory diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic system expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., hematopoietic cells and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

15

20

25

30

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to cytokine inducible inhibitor of signaling indicates that polypeptides and polynucleotides corresponding to Gene NO: 35 are useful for diagnosis and treatment of diseases of the hematopoietic system including autoimmune diseases, inflammatory diseases, infectious diseases and neoplasia. For example, administration of, or upregulation of this gene could by used to decrease the response of immune-system to lymphokines and cytokines.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 168 as residues: Arg-23 to His-30, Ala-35 to Gly-42.

FEATURES OF PROTEIN ENCODED BY GENE NO: 36

Gene NO: 36 is expressed primarily in infant brain and to a lesser extent in osteoclastoma, placenta, and a wide variety of other tissues.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., osteoclastoma, placenta, and tissue of the central nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 36 are useful for diagnosis and treatment of neurologic disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 169 as residues: Gln-31 to Ser-37, Ile-49 to Gly-54, Tyr-57 to Asp-67, Gln-141 to Pro-151, and Val-207 to Thr-219.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 37

Gene NO: 37 is expressed primarily in osteoclastoma stromal cells, dendritic cells, liver, and placenta.

WO 98/39446

5

10

15

25

30

35

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer, wound, pathological conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues or cell types (e.g., stromal cells, dendritic cells, liver, and placenta and, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 37 are useful for fundamental role in basic growth and development of human.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 170 as residues: Leu-32 to Thr-37 and Arg-48 to Pro-55.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 38

The translation product of Gene NO: 38 shares sequence homology with a yeast protein, Lpe10p, which may be involved in mRNA processing. (See Accession Nos. 2104457 and 1079682.) It is likely that an upstream signal sequence exists, other than the predicted sequence described in Table 1. Preferred polypeptide fragments comprise the open reading frame upstream from the predicted signal sequence, as well as polynucleotide fragments encoding these polypeptide fragments.

This gene is expressed primarily in skin, and to a lesser extent in embryonic tissues, and fetal liver.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, defects of the skin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skin, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., epidermis, liver, and embryanic tissues, and cancerous and wounded tissues) or bodily fluids (e.g., serum,

10

15

20

25

30

35

GENUARE AT

plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 38 are useful for diagnosis and treatment of defects of the skin.

FEATURES OF PROTEIN ENCODED BY GENE NO: 39

Gene NO: 39 is expressed primarily in Amygdala, activated monocytes, testis, and fetal liver. Moreover, this gene is mapped to chromosome 4. Thus, polynucleotides of the present invention can be used in linkage analysis as markers for chromosome 4.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, defects of the brain, immune system and testis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, immune system and testis, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., Amygdala, monocytes, testes, and liver and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 39 are useful for detecting defects of the brain, immune system and testis because of its abundance in these tissues.

FEATURES OF PROTEIN ENCODED BY GENE NO: 40

The translation product of Gene NO: 40 shares sequence homology with lymphoma 3-encoded protein (bcl-3) which is thought to contribute to leukemogenesis when abnormally expressed.

This gene is expressed primarily in Human Neutrophils, and to a lesser extent in Human Osteoclastoma Stromal Cells (unamplified), Hepatocellular Tumor, and Human Neutrophils, (Activated).

10

15

20

25

30

35

Principle (LLVV) Ophobaka,

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, chronic lymphocytic leukemia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., neutrophils, osteoclastoma, and kidney, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to lymphoma 3-encoded protein (bcl-3) indicates that polypeptides and polynucleotides corresponding to Gene NO: 40 are useful for treatment of lymphoma and related cancers.

FEATURES OF PROTEIN ENCODED BY GENE NO: 41

Gene NO: 41 is expressed primarily in ovary tumor, and to a lesser extent in endometrial stromal cells and fetal brain.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, ovarian or endometrial cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the female reproductive system and the developing central nervous system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., ovary, endometrium and brain, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

15

20

25

30

THE STEEL STATE OF SHARE

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 41 are useful for development of factors involved in ovarian or endometrial and general reproductive organ disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 174 as residues: Glu-22 to Trp-31, Asn-84 to Asp-90, and Ser-144 to Asp-151.

FEATURES OF PROTEIN ENCODED BY GENE NO: 42

The translation product of Gene 42 has sequence identity with a gene designated PTHrP(B). The PTHrP(B) polypeptide inhibits parathyroid hormone related peptide (PTHrP) activity.

This gene is expressed primarily in adult testis, and to a lesser extent in pituitary.

Therefore polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of male reproductive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., testes, and pituitary, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Furthermore, based in part on sequence identity with PTHrP(B), nucleic acids and polypeptides of the present invention may be used to diagnose or treat such conditions as hypercalcemia, osteoporosis, and disorders related to calcium metabolism.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 42 are useful for treatment of male reproductive disorders, hypercalcemia, osteoporosis, and other disorders related to calcium metabolism.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 175 as residues: Tyr-81 to Met-86, Gly-103 to Ser-108, Glu-127 to Pro-128, Pro-175 to Ser-180, Glu-196 to Lys-203, Pro-235 to Ser-241, and Ala-249 to Ser-264.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 43

The translation product of Gene NO: 43 shares sequence homology with brevican, which is thought to be important as a proteoglycan core protein of the

10

15

20

25

30

35

W. 2000244

aggrecan/versican family. The translation product of this gene may also contain a hyaluronan (HA)-binding region domain in frame with, but downstream of, the predicted open reading frame (Barta, et al., Biochem, J. 292:947-949 (1993)). The HA-binding domain, also termed the link domain, is found in proteins of vertebrates that are involved in the assembly of extracellular matrix, cell adhesion, and migration. It is about 100 amino acids in length. The structure has been shown to consist of two alpha helices and two antiparallel beta sheets arranged around a large hydrophobic core similar to that of C-type lectin. This domain typically contains four conserved cysteines involved in two disulfide bonds.

This gene is expressed primarily in early stage human brain and to a lesser extent in frontal cortex and epileptic tissues.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of disorders associated with, or observed during, neuronal development. Similarly, polypeptides and antibodies directed to these polypeptides are useful as immunological probes for differential identification of neuronal and associated tissues and cell types. For a number of disorders of the above tissues or cells, particularly for those of the nervous system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., brain and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to brevican indicates that polypeptides and polynucleotides corresponding to Gene NO: 43 are useful for neuronal regulation and signaling. The uses include directing or inhibiting axonal growth for the treatment of neuro-fibromatosis and in detection of glioses.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 176 as residues: Asp-28 to Arg-33 and Arg-126 to Arg-131.

FEATURES OF PROTEIN ENCODED BY GENE NO: 44

Gene NO: 44 is the human homolog of Notch-2 (Accession No. 477495) and mouse EGF repeat transmembrane protein (Accession No. 1336628), both genes are important in differentiation and development of an organism. The EGF repeat transmembrane protein is regulated by insulin like growth factor Type I receptor. These proteins are involved in cell-cell signaling and cell fate determination. Based on

10

15

20

25

30

35

homology, it is likely that this gene products also involved in cell differentiation and development. Although the predicted signal sequence is indicated in Table 1, it is likely that a second signal sequence is located further upstream. Moreover, further translated coding regions are likely found downstream from the disclosed sequence, which can easily be obtained using standard molecular biology techniques. A frameshift occurs somewhere around nucleotide 714, causing a frame shift in amino acid sequence from frame +2 to frame +3. However, using the homology of Notch-2 and EGF repeat transmembrane protein, the complete open reading frame can be elucidated. Preferred polynucleotide fragments comprise nucleotides 146-715, 281-715, and 714-965. Other preferred polypeptide fragments comprise the following EGF-like motifs: CRCASGFTGEDC (SEQ ID NO:260), CTCQVGFTGKEC (SEQ ID NO:261). CLNLPGSYQCQC (SEQ ID NO:262), CKCLTGFTGQKC (SEQ ID NO:263), and CQCLQGFTGQYC (SEQ ID NO:264).

Gene NO: 44 is expressed primarily in placenta and to a lesser extent in stromal and immune cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hemophelia and other blood disorders, central nervous system disorders, muscle disorders, and any other disorder resulting from abnormal development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, hematopoietic and vascular systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., placenta, stromal and immune cells and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to Notch-2 indicates that polypeptides and polynucleotides corresponding to Gene NO: 44 are useful for diagnosing and treating disorders relating to abnormal regulation of cell fate, induction, and differentiation of cells (e.g., cancer), epidermal growth factors, axonal pathfinding, and hematopoiesis.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 177 as residues: Gln-27 to Tyr-32, His-45 to Glu-55, Tyr-61 to Gly-77, Glu-99 to Ser-106, Ser-125 to Cys-131, and Thr-138 to Trp-144.

15

20

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 45

The translation product of this gene shares sequence homology with Laminin A which is thought to be important in the binding of epithelial cells to basement membrane and is associated with tumor invasion. Moreover, the translated protein is homologous to the *Drosophila* LAMA gene (Accession No. 1314864), a gene expressed in the first optic ganglion of *Drosophila*. Thus, it is likely that the gene product from this gene is involved in the development of the eye. Nucleotide fragments comprising nucleotides 822-1223, 212-475, 510-731, and 1677-1754 are preferred. Also preferred are the polypeptide fragments encoded by these polynucleotide fragments. It is likely that a frame shift occurs somewhere between nucleotides 475 to 510, shifting the open reading frame from +2 to +3. However, the open reading frame can be clarified using known molecular biology techniques.

This gene is expressed primarily in human testes tumor and to a lesser extent in placenta and activated monocytes.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, invasive cancers or tumors of the epithelium, as well as disorders relating to eye development. Similarly, polypeptides and antibodies directed to these polypeptides are useful as immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of neoplastic conditions, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., testes, placenta, and monocytes and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to Laminin A indicates that polypeptides and polynucleotides corresponding to Gene NO: 45 are useful for study and diagnosis of malignant or benign tumors, fibrotic disorders, and eye disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 178 as residues: Met-1 to Gly-8, Glu-32 to Ala-37, Met-113 to Asn-119, and Glu-139 to Gln-153.

FEATURES OF PROTEIN ENCODED BY GENE NO: 46

5

10

15

20

25

30

35

The translation product of Gene NO: 46 is novel and shares sequence homology with the product of the *Drosophila* tissue polarity gene frizzled. In vertebrates, it appears that there is a family of proteins that represent frizzled gene homologs. (See, e.g., Accession Nos. 1946343 and AFO17989.) The Drosophila frizzled protein is thought to transmit polarity signals across the plasma membrane of epidermal cells. The structure of frizzled proteins suggest that they may function as a G-protein-coupled receptor. The frizzled proteins are thought to represent receptors for Wnt gene products - secreted proteins that control tissue differentiation and the development of embryonic and adult structures. Inappropriate expression of Wnts has also been demonstrated to contribute to tumor formation. Moreover, mammalian secreted frizzled related proteins are thought to regulate apoptosis. (See Accession No. AFO17989.) The human homolog has also been recently cloned by other groups. (See Accession No. H2415415.) Thus, the protein encoded by this gene plays a role in mediating tissue differentiation, proliferation, tumorigenesis and apoptosis. Preferred polypeptide fragments lack the signal sequence as described in Table 1, as well as N-terminal and C-terminal deletions. Preferred polynucleotide fragments encode these polypeptide fragments.

Gene NO: 46 is expressed primarily in fetal tissues - particularly fetal lung - and adult cancers, most notably pancreas tumor and Hodgkin's lymphoma. Together, this distribution is consistent with expression in tissues undergoing active proliferation. The gene is also expressed to a lesser extent in other organs, including stomach, prostate, and thymus.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer (particularly pancreatic cancer and/or Hodgkin's lymphoma), as well as other forms of aberrant cell proliferation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and hyperproliferative disorders, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., fetal tissue, pancreas, and tissue of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,

10

15

20

25

30

35

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to frizzled indicates that polypeptides and polynucleotides corresponding to Gene NO: 46 are useful for influencing cell proliferation, differentiation, and apoptosis. The full-length protein or a truncated domain could potentially bind to and regulate the function of specific factors, such as Wnt proteins or other apoptotic genes, and thereby inhibit uncontrolled cellular proliferation. Expression of this protein within a cancer - such as via gene therapy or systemic administration - could effect a switch from proliferation to differentiation, thereby arresting the progression of the cancer.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 179 as residues: Pro-31 to Arg-37.

FEATURES OF PROTEIN ENCODED BY GENE NO: 47

The translation product of Gene NO: 47 shares sequence homology with members of the Rh/T2/S-glycoprotein family of ribonuclease-encoding genes. These ribonuclease proteins are found predominantly in fungi, plants, and bacteria and have been implicated in a number of functions, including phosphate-starvation response, self-incompatibility, and responses to wounding. A second group has recently cloned this same gene, calling it a ribonuclease 6 precursor. (See Accession No. 2209029.) This group also mapped the gene to chromosome 6, thus, the polynucleotides of the present invention can be used in linkage analysis as a marker for chromosome 6.

Gene NO: 47 is expressed primarily in hematopoietic cells and tissues, including macrophages, eosinophils, CD34 positive cells, T-cells, and spleen. It is also expressed to a lesser extent in brain and spinal cord.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, tumors of a hematopoietic origin, graft rejection, wounding, inflammation, and allergy. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., hematopoietic cells, and tissues and cells of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder.

relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to the Rh/T2/S-glycoprotein family of ribonuclease-encoding genes indicates that polypeptides and polynucleotides corresponding to Gene NO: 47 are useful as a cytotoxin that could be directed against specific cell types (e.g. cancer cells; HIV- infected cells), and that would be well tolerated by the human immune system.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 180 as residues: Ala-24 to Asp-30, Ile-51 to Tyr-61, Pro-69 to Ser-78, Pro-105 to Phe-110, Asn-129 to Phe-135, Pro-187 to Glu-192, Lys-205 to Gln-224, and Pro-250 to His-256.

FEATURES OF PROTEIN ENCODED BY GENE NO: 48

5

10

15

20

25

30

35

FREE P. L. W. GARMAGEA

The translation product of Gene NO: 48 shares sequence homology with dolichyl-phosphate glucosyltransferase, a transmembrane-bound enzyme of the endoplasmic reticulum which is thought to be important in N-linked glycosylation, by catalyzing the transfer of glucose from UDP-glucose to dolichyl phosphate. (See Accession No. 535141.) Based on homology, it is likely that this gene product also play a role similar in humans. Preferred polynucleotide fragments comprise nucleotides 132-959. Also preferred are the polypeptide fragments encoded by this nucleotide fragment.

Gene NO: 48 is expressed primarily in endothelial cells and to a lesser extent in hematopoietic cells and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, defects in proper N-linked glycosylation of proteins, such as Wiskott-Aldrich syndrome; tumors of an endothelial cell origin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular and hematopoietic systems, as well as brain, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell tpes (e.g., endothelial cells, hematopoietic cells, and brain, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

10

15

20

25

30

35

ARELONG JAN 1999 AR

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to dolichyl-phosphate glucosyltransferase indicates that polypeptides and polynucleotides corresponding to Gene NO: 48 are useful in diagnosing and treating defects in N-linked glycosylation pathways that contribute to disease conditions and/or pathologies.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 181 as residues: Lys-50 to Thr-55, Ser-73 to Arg-79, Glu-92 to Pro-99, Asp-110 to Ser-117, Gln-125 to Lys-131, Gly-179 to Asn-188, Ile-231 to Cys-236, and Glu-318 to Asn-324.

FEATURES OF PROTEIN ENCODED BY GENE NO: 49

Gene NO: 49 is expressed primarily in brain, most notably in the hypothalamus and amygdala. This gene is also mapped to chromosome X, and therefore, can be used in linkage analysis as a marker for chromosome X.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, tumors of a brain origin; neurodegenerative disorders, and sex-linked disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., brain and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.c., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 49 are useful for the diagnosis of tumors of a brain origin, and the treatment of neurodegenerative disorders, such as Parkinson's disease, and sex-linked disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 50

The translation product Gene NO: 50 shares sequence homology with canine phospholemman, a major plasma membrane substrate for cAMP-dependent protein kinases A and C. (See Accession No. M63934; see also Accession No. A40533.) In

WO 98/39446

5

10

15

20

25

30

35

fact, a group also recently cloned the human phospholemman gene, and mapped this gene to chromosome 19. (See Accession No.1916010.) Phospholemman is a type I integral membrane protein that gets phosphorylated in response to specific extracellular stimuli such as insulin and adrenalin. Phospholemman forms ion channels in the cell membrane and appears to regulate taurine transport, suggesting an involvement in cell volume regulation. It has been proposed that phospholemman is a member of a superfamily of membrane proteins, characterized by single transmembrane domains, which function in transmembrane ion flux. They are capable of linking signal transduction to the regulation of such cellular processes as the control of cell volume.

Gene No 50 is expressed primarily in fetal liver and to a lesser extent in adult brain and kidney, as well as other organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, insulin and/or adrenalin defects; diabetes; aberrant ion channel signaling; defective taurine transport; and defects in cell volume regulation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and/or immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., liver, brain, and kidney, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to phospholemman indicates that polypeptides and polynucleotides corresponding to Gene NO: 50 are useful for treatment of disorders involving the transport of ions and small molecules, in particular taurine. It could also be beneficial for control of pathologies or diseases wherein aberrancies in the control of cell volume are a distinguishing feature, due to the predicted role for phospholemman in the normal control of cell volume. It also may play a role in disorders involving abnormal circulating levels of insulin and/or adrenalin - along with other active secreted molecules - as revealed by its phosphorylation upon stimulation with insulin or adrenalin.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 183 as residues: Ala-20 to Gln-34, Arg-58 to Thr-79, and Leu-87 to Arg-92.

10

15

25

30

35

PNSTING CHWC 9F9944FAS

FEATURES OF PROTEIN ENCODED BY GENE NO: 52

Gene NO: 52 is expressed primarily in metastic melanoma and to a lesser extent in infant brain.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and cancer metastasis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., epidermis, and brain, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 52 are useful for diagnosis and treatment of melanoma.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 53

The translation product of Gene NO: 53 shares sequence homology with mucin which is thought to be important cell surface molecule. It also exhibits sequence identity with a calcium channel blocker of Agelenopsis aperta. In particular, with those calcium channel blockers which affect neuronal and muscle cells.

Gene NO: 53 is expressed primarily in prostate, endothelial cells, smooth muscle and fetal tissues and to a lesser extent in T cells and placenta.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, prostate cancer, immune disorders, angina, hypertension. cardiomyopathies, supraventricular arrhythmia, oesophogeal achalasia, premature labour, and Raynaud's disease. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues or cell types (e.g., prostrate, and tissue and cells of the immune system, and cancerous and wounded tissues) or

10

15

20

25

30

35

THE STATE OF THE S

bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to mucin indicates that polypeptides and polynucleotides corresponding to Gene NO: 53 are useful as a surface antigen for diagnosis of diseases such as prostate cancer and as tumor vaccine.

FEATURES OF PROTEIN ENCODED BY GENE NO: 54

Gene NO: 54 encodes a polypeptide which exhibits sequence identity with the rab receptor and VAMP-2 receptor proteins. (Martincic, et al., J. Biol. Chem. 272 (1997).)

Gene NO: 54 is expressed primarily in placenta, fetal liver, osteoclastoma and smooth muscle and to a lesser extent in T cell, fetal lung and colon cancer.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancers, osteoporosis and immuno-related diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, hematopoiesis system and bone system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., placenta, liver, osteoclastama, smooth muscle, T-cells, and lung, and colon, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 54 are useful for treating cancer, osteoporosis and immuno-disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 187 as residues: Pro-16 to Phe-21, Pro-24 to Arg-35, Arg-92 to Pro-98, Asn-143 to Lys-151, and Leu-169 to Ile-176.

10

15

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 55

Gene NO: 55 encodes a protein having sequence identity to the rat galanin receptor GALR2.

Gene NO: 55 is expressed primarily in ovarian cancer.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of ovarian cancer. Similarly, polypeptides and antibodies directed to those polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and reproductive system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., ovary, and tissues and cells of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. GALR2 antagonists can be used to treat obesity, bulimia, or Alzheimer's disease, while GALR2 agonists can be used to treat anorexia or pain, or to decrease nociception (claimed). Agonists and antagonists can also be used to treat numerous other disorders, including cognitive disorders, sensory disorders, motion sickness, convulsion/epilepsy, hypertension, diabetes, glaucoma, reproductive disorders, gastric and intestinal ulcers, inflammation, immune disorders, and anxiety.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 55 are useful for diagnosis and treatment of ovarian cancer.

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 56

As indicated in Table 1, the predicted signal sequence of Gene NO: 56 relates to an open reading frame that is homologous to the mouse major histocompatibility locus class III. (See Accession No. 2564953.) Any frame shift mutations that alter the correct open reading frame can easily be clarified using known molecular biology techniques. Moreover, in the opposite orientation, a second translated product is disclosed. This second translation product of this contig is identical in sequence to intracellular protein lysophosphatidic acid acyltransferase. The nucleotide and amino acid sequences of this translated product have since been published by Stamps and colleagues (Biochem. J. 326 (Pt 2), 455-461 (1997)), West and coworkers (DNA Cell Biol. 6, 691-701 (1997)), Rowan (GenBank Accession No. U89336), and Soyombo and Hofmann (GenBank Accession No. AF020544). This gene is thought to enhance cytokine

10

15

20

25

30

AREEDRAN TO THE THE

signaling response in cells. It is likely that a signal peptide is located upstream from this translated product. Preferred polypeptide fragments comprise the amino acid sequence: GLACWLAGVIFIDRKRTGDAISVMSEVAQTLLTQDVXVWVFPEGTRNHNGSML PFKRGAFHLAVQAQVPIVPIVMSSYQDFYCKKERRFTSGQCQVRVLPPVPTEGL TPDVPALADRVRHSMLHCF(SEQ ID NO: 265);

PSAKYFFKMAFYNGWILFLAVLAIPVCAVRGRNVENMKILRLMLLHIKYLYGI RVEVRGAHHFPPSQPYVVVSNHQSSLDLLGMMEVLPGRCVPIAKR (SEQ ID NO:266); TVFREISTD (SEQ ID NO:267); or LWAGSAGWPAG (SEQ ID NO: 268). Also provided are polynucleotide fragments encoding these polypeptide fragments.

Gene NO: 56 is expressed primarily in infant adrenal gland, hypothalamus, 7 week old embryonic tissue, fetal lung, osteoclastoma stromal cells, and to a lesser extent in a large number of additional tissues.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of developmental disorders and osteoclastoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s) in which it is highly expressed. For a number of disorders of the above tissues or cells, particularly during development or of the nervous or bone systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., adrenal, embryonic tissue, lung, and osteoclastomal stromal cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Further, expression of this protein can be used to alter the fatty acid composition of a given cell or membrane type.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 56 are useful for diagnosis and treatment of osteoclastoma and other bone and non-bone-related cancers, as well as for the diagnosis and treatment of developmental disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 189 as residues: Gly-29 to Gly-36 and Tyr-49 to Tyr-58.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 57

The translation product of Gene NO: 57 shares sequence homology with longevity-assurance protein-1. (See Accession No. g1123105.) Preferred

10

15

20

25

30

35

GENGLISH AT

ENSTROLL VW

polynucleotide fragments comprise nucleotides 6-125 and 118-432, as well as the polypeptides encoded by these polynucleotides. It is likely that a second signal sequence exists upstream from the predicted signal sequence in Table 1. Moreover, a frame shift likely occurs between nucleotides 118-125, which can be elucidated using standard molecular biology techniques.

Gene NO: 57 is expressed primarily in fetal liver, kidney, brain, thymus, and bone marrow.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immunological diseases and hyperproliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal liver, kidney, brain, thymus, and bone marrow expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., liver, kidney, brain, thymus, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to longevity-assurance protein suggest that Gene NO: 57 encodes a protein useful in increasing life span and in replacement therapy for those suffering from immune system disorders or hyperproliferative disorders caused by underexpression or overexpression of this gene.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 190 as residues: Val-29 to Arg-46 and Gly-50 to Gly-56.

FEATURES OF PROTEIN ENCODED BY GENE NO: 58

Domains of the Gene NO: 58 product are homologous to porcine surfactant protein-A receptor. (See Accession No. B48516.) The bovine gene binds surfactant protein-A receptor, modulating the secretion of alveolar surfactant. Based on this homology, the gene product encoded by this gene will likely have activity similar to the porcine gene. Preferred polynucleotide fragments comprise nucleotides 887-1039, as well as the polypeptide fragments encoded by this nucleotide fragment.

Gene NO: 58 is expressed primarily in brain and to a lesser extent in endothelial cells.

10

15

20

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the central nervous system including dimentia, stroke. neurological disorders, respiratory distress, and diseases affecting the endothelium including inflammatory diseases, restenosis, and vascular diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the placenta, liver, endothelial cells, prostate, thymus, and lung, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., brain, and endothelial cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology indicates that polypeptides and polynucleotides corresponding to Gene NO: 58 are useful for the diagnosis and /or treatment of diseases on the central nervous system, such as a factor that promote neuronal survival or protection, in the treatment of inflammatory disorders of the endothelium, or in disorders of the lung. In addition this protein may inhibit or promote angiogenesis and therefore is useful in the treatment of vascular disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 191 as residues: His-66 to Pro-80, Gly-139 to Ser-146 and Ser-262 to Pro-267.

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 59

The translation product of Gene NO: 59 is homologous to the rat hypertension-induced protein which is thought to be important in hypertension, and found expressed mainly in kidneys. (See Accession No. B61209.) Thus, it is likely that this gene product is involved in hypertension in humans. Preferred polypeptide fragments comprise the short chain dehydrogenase/reductase motif SILGIISVPLSIGYCASKHALRGFFNGLR (SEQ ID NO:269), as well as polynucleotides encoding this polypeptide fragment. Also preferred are polynucleotide fragments of 337-639, as well as the polypeptide fragments encoded by this polynucleotide fragment.

Gene NO: 59 is expressed primarily in liver, spleen, lung, brain, and prostate.

15

20

25

30

35

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cardiovascular, immunological, and renal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular, renal, and immune, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., liver, spleen, lung, brain, and prostrate, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to hypertension-induced protein indicates that polypeptides and polynucleotides corresponding to Gene NO: 59 are useful for treating hypertension.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 192 as residues: Gln-40 to Glu-45, Glu-96 to Glu-102, Asn-256 to Thr-266, and Asp-308 to Asp-317.

FEATURES OF PROTEIN ENCODED BY GENE NO: 60

Gene NO: 60 is expressed primarily in activated T-cell and jurkat cell and to a lesser extent in apoptic T-cell and CD34+ cell. It is likely that alternative open reading frames provide the full length amino acid sequence, which can be verified using standard molecular biology techniques.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, T lymphocyte related diseases or hematopoiesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., T-cells and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard

10

15

20

25

30

35

UK GLOWAT A.

100

gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 60 are useful for diagnosis or treatment of immune system disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 61

The translation product of Gene NO: 61, a vacuolar proton-ATPase, shares sequence homology with a *Caenorhabditis elegans* protein which is thought to be important in development. This protein may be a human secretory homologue that may also influence embryo development. Ludwig, J., also recently cloned this gene from chromaffin granules. (See, Accession No. 2584788.) Although Table 1 indicates the predicted signal peptide sequence, the translated product of this gene may in fact start with the upstream methionine, beginning with the amino acid sequence MAYHGLTV (SEQ ID NO:270). Thus, polypeptides comprising this upstream sequence, as well as N-terminus deletions, are also contemplated in the present invention.

Gene NO: 61 is expressed primarily in human placenta, liver, and Hodgkin's Lymphoma and to a lesser extent in bone marrow. Modest levels of expression were also observed in dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hyperproliferative disorders, defects in embryonic development, and diseases or disorders caused by defects in chromaffin granules. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly cancer, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., placenta, liver, lymph tissue, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to *Caenorhabditis elegans* indicates that polypeptides and polynucleotides corresponding to Gene NO: 61 are useful for diagnostic or therapeutic modalities for hyperproliferative disorders, embryonic development disorders, and chromaffin granules disorders.

10

15

20

25

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 62

The translation product of Gene NO: 62 shares sequence homology with the murine LAG3 gene which is thought to be important in the mediation of natural killer cell (NK cell) activity as previously determined by experiments in mice containing null mutations of LAG3. The similarity of this gene to the CD4 receptor may imply that the gene product may be a secreted, soluble receptor and immune mediator.

Gene NO: 62 is expressed primarily in human fetal heart, meningima, and to a lesser extent in tonsils. This gene also is expressed in the breast cancer cell line MDA 36.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, lymphomas, leukemias, breast cancer and any immune system dysfunction, including those dysfunctions which involve natural killer cell activities. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system or breast cancer, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., heart, meningima, and tonsils and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to the LAG3 gene (murine) indicates that the polynucleotides and polypeptides corresponding to Gene NO: 62 are useful for diagnostic and/or therapeutic modalities directed at abnormalities or disease states involving defective immune systems, preferably involving natural killer cell activity, as well as breast cancer.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 195 as residues: Pro-10 to Trp-17, Cys-58 to Pro-67, Thr-76 to Glu-85, and Arg-93 to Asn-101.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 63

The translation product of Gene NO: 63 shares sequence homology with a *Caenorhabditis elegans* alpha-collagen gene (Clg), which is thought to be important in

10

15

20

25

30

35

1000 1000

organism development, as well as other collagen genes. Thus, based on sequence homology, polypeptides of this gene are expected to have activity similar to collagen, including involvement in organ development.

Gene NO: 63 is expressed primarily in human B-Cell Lymphoma, and to a lesser extent in human pituitary tissue. This gene has also demonstrated expression in keratinocytes.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, B-Cell Lymphoma, other lymphomas, leukemias, and other cancers, as well as disorders related to development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., tissue and/or cells of the immune system, and pituitary, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to *Caenorhabditis elegans* alpha-collagen gene indicates that polypeptides and polynucleotides corresponding to Gene NO: 63 are useful for development of diagnostic and/or therapeutic modalities directed at the detection and/or treatment of cancer, specifically B-Cell Lymphomas, leukemias, or diseases related to development.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 196 as residues: Thr-22 to Arg-27 and Ser-29 to Thr-39.

FEATURES OF PROTEIN ENCODED BY GENE NO: 64

The translation product of Gene NO: 64 shares sequence homology with human extracellular molecule olfactomedin, which is thought to be important in the maintenance, growth, or differentiation of chemosensory cilia on the apical dendrites of olfactory neurons. Based on this sequence homology, it is likely that polypeptides of this gene have activity similar to the olfactomedin, particularly the differentiation or proliferation of neurons.

Gene NO: 64 is expressed primarily in fetal lung tissue.

10

15

20

35

() () (US S LL LA) ()

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the lung as well as neural development, particularly of the lung. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the pulmonary system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., lungs and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to the olfactomedin family indicates that polypeptides and polynucleotides corresponding to Gene NO: 64 are useful for the development of diagnostic and/or therapeutic modalities directed at detection and/or treatment of pulmonary disease states, e.g., cystic fibrosis.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 197 as residues: Gly-17 to Gln-23, Gln-45 to Arg-50, Arg-56 to Lys-61, Glu-70 to Leu-76, Asp-88 to Glu-93, Pro-117 to Met-131, Asp-161 to Glu-167, Arg-224 to Asn-237, Asp-302 to Trp-312, Pro-315 to Asn-320, and Thr-337 to Ser-341.

FEATURES OF PROTEIN ENCODED BY GENE NO: 65

The translation product of Gene NO: 65 shares sequence homology with

Saccharomyces cerevisiae hypothetical protein YKL166 (Accession No. gi/687880)

which is thought to be important in secretory and/or vesicular transport mechanisms.

Based on this homology, it is likely that the gene product would have similar activity to YKL166, particularly secretory or transport mechanisms. Preferred polypeptide fragments of this gene include those fragments starting with the amino acid sequence

ISAARV (SEQ ID NO:271). Other polypeptide fragments include the former fragment, which ends with the amino acid sequence PDVSEFMTRLF (SEQ ID NO:272). Further preferred fragments include those polypeptide fragments comprising the amino acid sequence FDPVRVDITSKGKMRAR (SEQ ID NO:273). Also preferred are polypeptide fragments having exogenous signal sequences fused to the polypeptide.

Gene No 65 is expressed primarily in placenta, testis, osteoclastoma and to a lesser extent in adrenal gland.

10

15

20

25

30

35

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and/or diseases involving defects in protein secretion. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, cartilage and bone, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., placenta, testis, adrenal gland, and osteoclastoma, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to the yeast YKL1GG protein indicates that polypeptides and polynucleotides corresponding to Gene NO: 65 are useful for the development of therapeutic and/or diagnostic modalities targeted at cancer or secretory anomalies, such as genetically caused secretory diseases.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 198 as residues: Ser-18 to Ser-29 and Lys-53 to Arg-74.

FEATURES OF PROTEIN ENCODED BY GENE NO: 66

The translation product of Gene NO: 66 shares sequence homology with the human papilloma virus (HPV) E5 ORF region which is thought to be important as a secreted growth factor. Although this is described as a viral gene product, it is believed to have several cellular secretory homologues. Therefore, based on the sequence similarity between the HPV E5 ORF and the translated product of this gene, this gene product is likely to have activity similar to HPV E5 ORF.

Gene NO: 66 is expressed primarily in activated T-Cells, monocytes, cerebellum and to a lesser extent in infant brain.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and/or human papilloma virus infection. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of

10

15

20

25

30

35

this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., brain, lymph tissue, monocytes, and T-cells and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Moreover, polynucleotides of this gene have been mapped to chromosome 1. Therefore, polynucleotides of the present invention can be used in linkage analysis as a marker for chromosome 1.

The tissue distribution and homology to human papilloma virus E5 region indicates that polypeptides and polynucleotides corresponding to Gene NO: 66 are useful for development of diagnostic and/or therapeutic modalities directed at the diagnosis and/or treatment of cancer and/or human papilloma virus infection (HPV).

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 199 as residues: Asn-31 to Arg-36 and Leu-102 to Ser-112.

FEATURES OF PROTEIN ENCODED BY GENE NO: 67

The translation product of Gene NO: 67 shares sequence homology with the 8hs20 protein precursor [Mus musculus] which is thought to be important in B-Cell mu chain assembly. (See, Accession No. PID/d1002996; Shiraswa, T., EMBO. J. 12(5):1827-1834 (1993).) A polypeptide fragment starting at amino acid 53 is preferred, as well as 1-20 amino acid N-terminus and/or C-terminus deletions. Based on the sequence similarity between 8hs20 protein and the translation product of this gene, the two polypeptides are expected to share certain biological activities, particularly immunologic activities.

Gene NO: 67 is expressed primarily in human B-cells and to a lesser extent in Hodgkin's Lymphoma. It is also likely that the polypeptide will be expressed in B-cell specific cells, bone marrow, and spleen, as is observed with 8hs20.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, Hodgkin's Lymphoma, Common Variable Immunodeficiency, and/or other B-cell lymphomas. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., bone

10

15

20

25

30

35

marrow, spleen, lymph tissue, and B-cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to 8hs20 protein precursor [Mus musculus], indicates that polypeptides and polynucleotides corresponding to Gene NO: 67 are useful for therapeutic and/or diagnostic purposes, targeting Hodgkin's Lymphoma, B-cell lymphomas, Common Variable Immunodeficiency, or other immune disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 200 as residues: Asp-51 to Trp-56, Arg-72 to Asp-85, and Gln-106 to Asp-112.

FEATURES OF PROTEIN ENCODED BY GENE NO: 68

Gene NO: 68 is expressed primarily in fetal liver/spleen, rhabdomyosarcoma, and to a lesser extent in 9 week-old early stage human embryo and bone marrow.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, rhabdomyosarcoma and other cancers, hematopoietic disorders, and immune dysfunction. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., embryonic tissue, striated muscle, liver, spleen, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that theprotein product of Gene NO: 68 is useful for diagnostic and/or therapeutic purposes directed to cancer, preferably rhabdomyosarcoma. Enhanced expression of this gene in fetal liver, spleen, and bone marrow indicates that this gene plays an active role in hematopoiesis. Polypeptides or polynucleotides of the present invention may therefore help modulate survival, proliferation, and/or differentiation of various hematopoietic lineages, including the hematopoietic stem cell. Thus, polynucleotides or polypeptides can be used treat

various hematopoietic disorders and influence the development and differentiation of blood cell lineages, including hematopoeitic stem cell expansion. The polypeptide does contain a thioredoxin family active site at amino acids 64-82. Polypeptides comprising this thioredoxin active site are contemplated.

5

10

15

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 69

Gene NO: 69 is expressed primarily in liver and kidney and to a lesser extent in macrophages, uterus, placenta, and testes.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, renal disorders, neoplasms (e.g., soft tissue cancer, hepatacellular tumors), immune disorders, endocrine imbalances, and reproductive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatic, urogenital, immune, and reproductive systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., liver, kidney, uterus, placenta, testes, and macrophages and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder. relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 69 are useful for diagnosis and treatment of disorders in the hepatic, urogenital, immune, and reproductive systems.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 202 as residues: Arg-41 to Ser-50, Glu-138 to Asn-148, Ser-155 to Arg-172, Pro-219 to Glu-228.

30

35

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 70

Gene NO: 70 is expressed primarily in the immune system, including macrophages, T-cells, and dendritic cells and to a lesser extent in fetal tissue.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, inflammatory diseases, lymph node disorders, fetal

10

15

20

25

30

ENCOUNT ON A CONTRACTOR

development, and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and certain cell types (e.g., macrophages, T-cells, dendritic cells, and fetal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. There is some evidence that the polynucleotide is mapped to chromosome 19. Thus, the polynucleotide can be a marker for genetic analysis for chromosome 19.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 70 are useful for treatment, prophylaxis, and diagnosis of immune and autoimmune diseases, such as lupus, transplant rejection, allergic reactions, arthritis, asthma, immunodeficiency diseases, leukemia, and AIDS. The polypeptides or polynucleotides of the present invention are also useful in the treatment. prophylaxis, and detection of thymus disorders, such as Graves Disease, lymphocytic thyroiditis, hyperthyroidism, and hypothyroidism. The expression observed predominantly in hematopoietic cells also indicates that the polynucleotides or polypeptides are important in treating and/or detecting hematopoietic disorders, such as graft versus host reaction, graft versus host disease, transplant rejection, myelogenous leukemia, bone marrow fibrosis, and myeloproliferative disease. The polypeptides or polynucleotides are also useful to enhance or protect proliferation, differentiation, and functional activation of hematopoietic progenitor cells (e.g., bone marrow cells), useful in treating cancer patients undergoing chemotherapy or patients undergoing bone marrow transplantation. The polypeptides or polynucleotides are also useful to increase the proliferation of peripheral blood leukocytes, which can be used in the combat of a range of hematopoietic disorders, including immunodeficiency diseases, leukemia, and septicemia.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 203 as residues: Thr-21 to Ser-27, Pro-33 to Ser-38, and Arg-73 to Lys-84.

Last AA of OR F	466	221	34	155	232	42
Predicted First AA of Secreted Portion	29	29	30	9 ٤	21	Z.
Last AA of Sig Pep	28	28	56	35	20	31
irst AA of Sig	_	-	—	_	-	_
AA SEQ ID NO: Y	13.1	१३६।	204	136	137	205
S'NT AA F of AA F First SEQ AA of ID Signal NO: Signal NO: Y	54	68	01	173	202	861
of of Start	54	30	<u>C</u>	173	202	
3' NT of Clone Seq.	1658	844	434	676	1343	1309
S' NT 3' NT of of Soft Clone Seq. Seq.	25	_	_	134	727	741
Tota NT Seq	1739	844	795	776	1376	1324
SEQ ID NO:	=	12	<u>~</u>	<u>~</u> ,	<u> </u>	85
Vector	p.Sport1	pCMVSport 3.0	pCMVSport 3.0	Uni-ZAP XR	pBluescript	pBluescript
ATCC Deposit No: Z and Date	97901 02/26/97 209047 05/15/97	97898 02/26/97 209044 05/15/97	97898 02/26/97 209044 05/15/97	97899 02/26/97 209045 05/15/97	97900 02/26/97 209046 05/15/97	97900 02/26/97 209046 05/15/97
cDNA Clone ID	HGCMD20	HLDBG33	HLDBG33	HTGEW86	HKCSR70	HKCSR70
Gene No.	_	2	2	κ,	- 1	4

d Last AA OR OR	84	40	72	9 <u>/</u> £	207	42
Predicte First AA of Secreted Portion	35	34	81	27	56	22
Last AA of Sig Pep	74	33	17	2k	28	21
rirst AA of Sig Pep	_	_	_	_		
AA SEQ NO:	206	138	139	140	207	141
F First SEQ AA of ID Signal NO:	51	143	95	45	5	157
5' NT 3' NT of of 5' NT of Clone Clone of Start Seq. Seq. Start	51	143	98	45	5	157
3' NT of Clone Seq.	1484	502	425	1298	1271	384
5' NT of Clone Seq.	_		_	_		87
Total NT Seq.	1494	502	425	1316	1285	436
SEQ NO:	83	5.1	16	17	84	<u>«</u>
Vector	Uni-ZAP XR	Uni-ZAP XR	pBluescript	Uni-ZAP XR	Uni-ZAP XR	pBluescript
ATCC Deposit No: Z and Date	209010 04/28/97 209085 05/29/97	97897 02/26/97 209043 05/15/97	97897 02/26/97 209043 05/15/97	97897 02/26/97 209043 05/15/97	97897 02/26/97 209043 05/15/97	97897 02/26/97 209043 05/15/97
cDNA Clone ID	HETBI87	HTEAU17	HBMCY91	HSSGE07	HSSGE07	HBMBX59
Gene No.	4	8	9	7		<u></u>

Last AA of OR F	40	69	182	23	482	12
Predicted First AA Last of AA Secreted of Portion OR	20	32	<u>15</u>	21	31	
First Last AA AA of of Sig Sig Pep Pep	19	31	30	50	30	
First AA of Sig Pep	_		_	_		_
AA SEQ ID NO:	1.12	143	17 17	208	209	210
Signal NO: Pep Y	£.7	147	157	991	157	1137
S' NT 3' NT of of S' NT Clone Clone of NT Seq. Seq. Start Seq.	23	147	157	991	157	
3' NT of Clone Seq.	503	358	1926	394	1925	1298
5. NT of Clone Seq.	_	_	573	_	573	30
Total NT Seq.	503	358	9761	394	1925	1818
NT SEQ ID NO:	6	20	<u></u>	8.5	98	87
Vector	Uni-ZAP XR					
ATCC Deposit No: Z and Date	97897 02/26/97 209043 05/15/97	97897 02/26/97 209043 05/15/97	97898 02/26/97 209044 05/15/97	97898 02/26/97 209044 05/15/97	97898 02/26/97 209044 05/15/97	97898 02/26/97 209044 05/15/97
cDNA Clone ID	HNGIT22	HERAD57	HCEN140	HCENJ40	HCENJ40	HCEN:140
Gene No.	6	10		_		_

Last AA of OR	225	44	6	131	54	91
Predicted First AA Last of AA Secreted of Portion OR	31	40	6	31	38	₹.
Last AA of Sig Pep	30	36	<u>×</u>	30	37	30
irst AA of Sig	_		_	-	,	_
SEQ NO:	145	146	211	147	212	148
of AA F of First SEQ / AA of ID Signal NO: 9	80	181	215		513	7.1
S NT 3' NT of S' NT of Of S' NT of Clone Clone of A Start Seq. Start S	80	181	512	_	513	77
of of Clone Seq.	557	694	539	962	855	653
of Clone Seq.	64	_	_	405	300	205
Total NT Seq.	1224	694	530	961	855	299
SEQ × SOC:	22	23	∝ ∝	24	89	52
Vector	Uni-ZAP XR					
ATCC Deposit No: Z and Date	97898 02/26/97 209044 05/15/97	97898 02/26/97 209044 05/15/97	97898 02/26/97 209044 05/15/97	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97
cDNA Clone ID	HCSRA90	HBJFC03	HBJFC03	HSNBL.85	HSNBL85	HTERY26
Gene No.	12	13	<u></u>	4	4	<u>~</u> .

Last AA of OR F	34	164	520	138	126	57
Predicted First AA of Secreted Portion	Z£	6	r:	<u>.</u> .	28	0٤
Last AA of Sig Pep	31	<u>«</u>	22	30	27	29
rst A of ig			_		_	-
AA SEQ NO:	213	149	214	150	917	151
Signal NO: Sep A Find Pierst SEQ A Find Signal NO: Signal NO: Sep A Find Pep	275	& &	70	97	100	169
of Of Start		∝ ∝	70	97	001	169
3' NT of Clone Scq.	625	1105	6001	1017	943	391
Seq. Seq.	861	40	19	_	-	_
Total NT Seq.	628	1105	1053	1017	2492	391
SEQ NO DO	06	26	<u> </u>	27	£6	28
Vector	Uni-ZAP XR	Uni-ZAP XR	Hpi-ZAP XR	pBluescript	pBluescript	Uni-ZAP XR
ATCC Deposit No: Z and Date	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97
cDNA Clone ID	HTEBY26	HMABH07	HMAB1107	HSKNY94	HSKNY94	HMCDA67
Gene No.	15	91	91	17	17	<u>~</u>

Last AA of OR F	47	46	4	40	-	105
Predicted First AA Last of AA Secreted of Portion OR	45	47	29	34	25	48
Last AA of Sig Pep	44	46	28	ι.	24	47
First AA of Sig Pep	_			_		-
AA SEQ ID NO: Y	152	217	153	218	154	155
of AA First Last P First SEQ AA AA F AA of ID of of of Signal NO: Sig	109	1868	47	699	403	49
Seq. Seq. Start Seq. Codon	109	1868	47	699	403	49
3' NT of Clone Seq.	1139	2847	370	1000	702	518
S' NT of Clone Seq.	9	3058 1795 2847	_	664	-	
Total NT Seq.	1139	3058	465	1099	702	1142
NT SEQ ID NO:	29	94	30	95	۲,	32
Vector	Uni-ZAP XR	Uni-ZAP XR	pSport1	pSport1	Uni-ZAP XR	Uni-ZAP XR
ATCC Deposit No: Z and Date	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97
cDNA Clone ID	HOSFF45	HOSFF45	HMJAA51	HMJAA51	HTEBEOS	HTEAL31
Gene No.	19	19	20	20	21	22

Alberta : We - Greensh

Last AA of OR F	104	82	% %	25.	91	74
First Last Predicted AA AA First AA Last of of of AA Sig Sig Secreted of Pep Portion OR F	48	28	82	23		26
Last AA of Sig Pep	47	27	27	22		25
First AA of Sig Pep		-		_		_
AA SEQ NO:	219	981	220	157	221	158
5' NT of First AA of Signal Pep	32	48	68	30	507	40
S' NT 3' NT of AA Fir. of of 5' NT First SEQ AA tal Clone Clone of AA of ID of T Seq. Seq. Start Signal NO: Signal Godon Pep Y Pe	£.	48	68	30	507	40
3' NT of Clone Seq.	422	928	593	773	1253	453
5° NT of Clone Seq.	23	_	72	_	507	
C _Z S	1580	928	829	£22	1253	453
SEQ SEQ NO:	96	33	97	3.4	86	35
Vector	Uni-ZAP XR	pBluescript	pBluescript	pBluescript	pBluescript	Uni-ZAP XR
ATCC Deposit No: Z and Date	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97	97899 02/26/97 209045 05/15/97
cDNA Clone ID	HTEAL31	HBMCT32	HBMCT32	HSKXE91	HSKXE91	HPWTB39
Gene No.	22	23	23	2.4	24	25

Last AA of OR F	0 <u>%</u>	-38 -	137	177	49	71
Predicted First AA of Secreted Portion	۲. ک	20	24	C 1	27	<u>C</u> 2
Last AA of Sig Pep	24	61	23	12	26	
rst A of ig ep	_	_	_	_	_	
AA SEQ ID NO: Y	159	160	222	161	223	162
5° NT of First AA of Signal Pep	25	_	_		17	
5' NT 3' NT of AA Fi of AA Fi of Of S' NT First SEQ A Of Seq. Seq. Start Signal NO: Seq. Codon Pep Y P	52		7		17	
3' NT of Clone Seq.	459	509	447	598	611	454
5' NT of Clone Seq.	_	_	_		37	_
Total NT Seq.	450	509	447	\$68	611	454
NT SEQ ID NO:	9દ	37	66	&, &,	001	6£
Vector	Uni-ZAP XR	pSport1	pSport1	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR
ATCC Deposit No: Z and Date	97899 02/26/97 209045 05/15/97	97900 02/26/97 209046 05/15/97	97900 02/26/97 209046 05/15/97	97900 02/26/97 209046 05/15/97	97900 02/26/97 209046 05/15/97	97900 02/26/97 209046 05/15/97
cDNA Clone ID	HTLEV12	HSPAF93	HSPAF93	HHFGL.62	HHFGL62	HCEIU14
Gene No.	56	27	27	28	28	29

PROPERTY OF THE STATE OF

Last AA of OR F	14	99	154	154	6	<u>103</u>
Predicted First AA of Secreted Portion		61	<u></u>	32		19
AA First Last SEQ AA AA Of of of NO: Sig Sig Y Pep Pep		18	30	31		18
irst AA of Sig				-		_
AA SEQ ID NO: Y	224	163	164	225	226	165
5' NT of First AA of Signal Pep	237	223	213	611	138	611
of of Start	152	223	213	119	138	119
3' NT of Clone Seq.	609	376	2471	1721	1777	2659
5' NT of Olone (Clone (Seq.	176		141	47	96	1172
Total NT Seq.	609	425	2471	1770	1832	2650
NT SEQ ID NO:	101	40	41	102	103	42
Vector	Uni ZAP XR	Uni-ZAP XR				
ATCC Deposit No: Z and Date	97900 02/26/97 209046 05/15/97	97900 02/26/97 209046 05/15/97	97900 02/26/97 209046 05/15/97	97900 02/26/97 209046 05/15/97	97900 02/26/97 209046 05/15/97	97900 02/26/97 209046 05/15/97
cDNA Clone ID	HCEIU14	HEBDA39	HTHBA79	IITHBA79	HTHBA79	HAGBB70
Gene No.	29	30	<u></u>	« ,	31	35

Last AA of OR F	61	<u>@</u>	92	93	93	57	36
Predicted First AA of Secreted Portion		21	24	24	22	2	24
Last AA of Sig Pep		50	23	23	21	30	£.;
irst AA of of Sig	_	_	_		_		_
AA SEQ ID NO: Y	227	166	167	228	229	168	230
of AA F First SEQ AA of ID Signal NO: Signal	1134	299	<u>C</u>	272	168	1437	080
Seq. Seq. Start Seq. Start Seq.	1134	599	<u>C</u>	272	891	1437	686
3' NT of Clone Seq.	2237	1580	717	1023	6991	2378	1892
5' NT of Clone Seq.	878	001	61	_		1337	1068
Total NT Seq.	7577	1635	780	1822	106 1712	2378	1969
NT SEQ ID NO:	104	43	44	105	901	45	107
Vector	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	pSport1	pBluescript	Uni-ZAP XR	Uni-ZAP XR
ATCC Deposit No: Z and Date	97900 02/26/97 209046 05/15/97	97900 02/26/97 209046 05/15/97	97900 02/26/97 209046 05/15/97	209236 09/04/97	209084 05/29/97	97900 02/26/97 209046 05/15/97	97900 02/26/97 209046 05/15/97
cDNA Clone ID	HAGBB70	HETDG84	HTEGA81		HKMLK44	HTXAK60	HTXAK60
Gene No.	Z£	33	34	34	34	35	35

Last AA of OR F	231	08	71	64	74	333
First Last Predicted AA A First AA Last of of of AA Sig Sig Secreted of Pep Pep Portion OR	31	30	.	24	23	2
Last AA of Sig Pep	30	56	30	23	22	_
First AA of Sig Pep	_	-			-	-
AA SEQ ID NO: Y	169	231	170	171	172	173
Signal NO:	129	001	€.	167	364	2
s: NT of Start Codon	129	100	<u>د.</u>	167	364	2
3' NT of Clone Seq.	1772	1734	1107	764	1258	1184
5. NT of Clone Seq.	69	59	70	167	131	
Total NT Seq.	1772	1734	1107	805	1408	1813
SEQ 1D NO:	46	108	47	48	49	05
Vector	Uni-ZAP XR	Uni-ZAP XR	pBluescript	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR
ATCC Deposit No: Z and Date	97901 02/26/97 209047 05/15/97	97901 02/26/97 209047 05/15/97	97901 02/26/97 209047 05/15/97	97901 02/26/97 209047 05/15/97	97901 02/26/97 209047 05/15/97	97901 02/26/97 209047 05/15/97
cDNA Clone ID	HMHBN40	HMHBN40	HFVGS85	HERAH81	HMSEU04	HNEDIS7
Gene No.	9£	36	37	&.	39	40

Last AA of OR	561	300	264	312	137	47
Predicted First AA of Secreted Portion	21	23	26	30	23	34
	20	22	2.5	29	22	33
	_	-		_	-	
SEQ NO: NO:	174	232	571	233	176	234
5° NT of First AA of Signal Pep	142	89	158	14	191	999
T of AA First SEQ AA of ID of Start Signal NO: Signal AY Pep	142	89	158	41	191	
3' NT of Clone Seq.	2070	1957	1426	1311	1720	1962
S NT 3' NT of of S	74	51		08	_	299
Total NT Seq.	2070	109 2003	1426	1320	1720	1962
SEQ ID NO:	51	601	52	110	۲,	
Vector	pSport1	pSport I	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR
ATCC Deposit No: Z and Date	97901 02/26/97 209047 05/15/97	97901 02/26/97 209047 05/15/97	97901 02/26/97 209047 05/15/97	97901 02/26/97 209047 05/15/97	97901 02/26/97 209047 05/15/97	97901 02/26/97 209047 05/15/97
cDNA Clone ID	HNTME13	HNTME13	HSXB125	IISXB125	HSXCK41	HSXCK41
Gene No.	41	41	42	42	43	43

d Last A A GR	178	23	154	315	294	295
Predicted First AA 1 of Secreted Portion (56	£.1	32	74.	٠, 2	25
Last AA of Sig Pep	52	£.7	[,	36.	57	24
irst AA of Sig	_	_	_		_	_
AA SEQ ID NO:	177	235	178	236	179	237
Signal NO: 8	218	225	611	0%	124	165
of of Start	218		611	V8	124	165
3' NT of Clone Seq.	1107	1087	1903	1832	1838	1960
S' NT 3' NT of of Soficial Seq. Seq.	_	30			133	06
Tota NT Seq.	1117	1785	1903	1842	1869	1960
SEQ NO:	54	112	55	-13	56	114
Vector	Uni-ZAP XR					
ATCC Deposit No: Z and Date	97902 02/26/97 209048 05/15/97	97902 02/26/97 209048 05/15/97	97902 02/26/97 209048 05/15/97	97902 02/26/97 209048 05/15/97	97902 02/26/97 209048 05/15/97	97902 02/26/97 209048 05/15/97
cDNA Clone ID	IIE8C126	HE8C.126	HTTDS54	HTTDS54	HICHDY31	HLJIDY31
Gene No.	44	14	45	45	46	46

Last AA of OR F	255	323	46	92	91	42
Predicted First AA of Secreted Portion	27	61	38	£9	23	30
Last AA of Sig Pep	26	18	34	79	22	29
irst of of Sig	_	_			-	-
SEQ NO:	180	<u>8</u>	182	<u>83</u>	238	185
of AA F of First SEQ / AA of ID Signal NO: 8	352	12	172	40	73	308
S' NT 3' NT of S' NT of Clone Clone of Start Seq. Seq. Scoton	352	12	172	40	73	308
3' NT of Clone Seq.	1010	557	304	501	536	595
5' NT of Clone Seq.	320	33		_	73	_
Total NT Seq.	1259	1186	428	105	536	595
SEQ NO:	75	58	59	09	115	62
Vector	Uni-ZAP XR	pSport1				
ATCC Deposit No: Z and Date	97902 02/26/97 209048 05/15/97	97902 02/26/97 209048 05/15/97	97902 02/26/97 209048 05/15/97	97902 02/26/97 209048 05/15/97	97902 02/26/97 209048 05/15/97	97903 02/26/97 209049 05/15/97
cDNA Clone ID	HMCBP63	HEMGE83	HHSDC22	HHSDZ57	HHSDZ57	HMMAB12
Gепе No.	47	48	49	50	20	52

Last
AA
of
OR
F

28

57

187

122

Predicted First AA of Secreted Portion	7.2	40	92	31	54	<i>L</i> 2
Last AA of Sig Pep	56	39	25	30	£73	26
First AA of Sig Pep	-	_			_	-
AA SEQ ID NO: Y	241	186	242	187	243	<u>~</u>
S' NT of First AA of Signal Pep	198	176	317	30	296	_
Soft Seq. Seq. Seq. Seq. Seq. Seq. Seq. Seq.	108	176	317	30	596	_
3. NT of Clone Seq.	453	1436	1957	2033	2134	440
5' NT of Clone Seq.		40	211		110	_
Total NT Seq.	153	1478	2016	2011	2136	4.10
SEQ SEQ NT NO:	<u>~</u> 8.	63	611	64	120	65
Vector	pSport1	Uni-ZAP XR	Uni-ZAP XR	Uni ZAP XR	Uni-ZAP XR	Uni-ZAP XR
ATCC Deposit No: Z and Date	97903 02/26/97 209049 05/15/97	97903 02/26/97 209049 05/15/97	97903 02/26/97 209049 05/15/97	97903 02/26/97 209049 05/15/97	97903 02/26/97 209049 05/15/97	97903 02/26/97 209049 05/15/97
cDNA Clone ID	HMMABIZ	HSKDW02	HSKDW02	HETGL41	HETGL41	HODAZ50
Genc No.	25	53	53	5.1	54	λ.

Last AA of OR	72	83	75	48	310	3,38
Predicted First AA Last of AA Secreted of Portion OR	=	31	27	28	31	<u>-</u>
AAA of of Sig	C	30	26	27	<u>30</u>	30
First AA of Sig Pep		_		-	-	_
AA SEQ ID NO:	244	189	U61	245	161	246
of AA First I Frirst SEQ AA AA of ID of Signal NO: Sig	-	341	331	367	57	08
S' NT 3' NT of of S' NT Clone Clone of Seq. Seq. Start Codon		341	33		57	08
3' NT of Clone Seq.	219	1478	1535	1678	1244	1211
5' NT of Clone Seq.	_	349		239	402	_
Total NT Seq.	219	3301	रेहेंडे।	1686	1244	1211
SEQ SEQ NO:	121	99	67	122	89	123
Vector	Uni-ZAP XR					
ATCC Deposit No: Z and Date	97903 02/26/97 209049 05/15/97	97903 02/26/97 209049 05/15/97	97903 02/26/97 209049 05/15/97	97903 02/26/97 209049 05/15/97	97903 02/26/97 209049 05/15/97	97903 02/26/97 209049 05/15/97
cDNA Clone ID	HODAZ50	HSDGE59	HE6FS13	HE6ES13	HSSEP68	HSSEP68
Gene No.	55	95	23	57	δ ,	φ.

Last AA of OR F	17	317	3,38	52	-4	<u>-</u>
		ر.,	ι,	۷,		
First Last Predicted AA AA First AA of of of Sig Sig Secreted Pep Pep Portion		29	22	18	50	4 3
Last AA of Sig Pep		28	21	30	28	42
First AA of Sig Pep	-	_	_		_	_
AA SEQ ID NO:	247	701 201	248	193	194	195
F. NT First SEQ A Post Start Signal NO:	501	70	70	985	187	118
5' NT of Start Codon	501	70	70	536	187	118
S' NT 3' NT of of Clone Clone Seq. Seq.	1526	1278	1088	1031	855	1274
S NT 3 NT of of S	402	_	3	498	178	28
Total NT Seq.	1804	<u> 2</u> 021	1282	1031	855	1274
X SEQ NO DEQ	124	69	125	70	17	72
Vector	Uni-ZAP XR	Uni ZAP XR	Uni-ZAP XR	pBluescript SK-	Uni-ZAP XR	Uni-ZAP XR
ATCC Deposit No: Z and Date	97903 02/26/97 209049 05/15/97	97903 02/26/97 209049 05/15/97	97903 02/26/97 209049 05/15/97	97903 02/26/97 209049 05/15/97	97904 02/26/97 209050 05/15/97	97904 02/26/97 209050 05/15/97
cDNA Clone ID	HSSEP68	HRDEV41	IIRDEV41	HILCJ01	HSATP28	HIIFGL41
Gene No.	28	65	65	09	19	62

HAND TO WO SERVERER,

Last AA of OR F	ν,	4	78	354	353	73
	\$6	77	7	κ.΄	ξ.	7
Predicted First AA of Secreted Portion	40	61	2 1	22	24	61
First Last AA AA of of Sig Sig Pep Pep	36	<u>~</u>	20	21	23	<u>«</u>
First AA of Sig Pep	_		-	_		_
ASEQ Yö: BÖ:	249	961	250	197	251	198
of AA F First SEQ AA of ID Signal NO: 8	133	173	174	112	8	531
Seq. Seq. Seq. Start Si	133	173	174	112	87	531
3' NT of Clone Seq.	1237	889	737	1890	1829	1133
5° NT of Clone Seq.	<u></u> 88 .		_	_	_	408
Total NT Seq.	1296	889	737	1890	1925	1133
NT SEQ ID NO:	126	73	127	74	128	75
Vector	Uni-ZAP XR					
ATCC Deposit No: Z and Date	97904 02/26/97 209050 05/15/97	97904 02/26/97 209050 05/15/97	97904 02/26/97 209050 05/15/97	97904 02/26/97 209050 05/15/97	97904 02/26/97 209050 05/15/97	97904 02/26/97 209050 05/15/97
cDNA Clone ID	HHFGL41	HBJEM49	HBJEM49	HSLD195	HSLDJ95	HSREG44
Gene No.	62	63	63	64	64	59

4 44504 49 75 4 7 1 1 1 1 1 1

Last AA of OR F	112	108	122	2 1 °.	44	314	235
Predicted First AA L of Secreted Portion	70	40	24	t		28 3	
Last AA of Sig Pep	69	30	23	23	50	27	9
AA of of Sig Pep	-			_	-	I	
A SEQ YÖÖ:	100	252	500	102	253	254	202
Signal NO:		<u> </u>	51	2.5	701	25	95
5° NT of Start Codon	_	2133	<u>~</u> ,	٠ <u>٠</u>	701	25	95
S' NT 3' NT of of Clone Clone Seq. Seq.	585	2713	577	1935	101	1929	1097
S NT 3' NT of of Clone Clone Seq. Seq.	_	2023		1458	479	_	601
Total NT Seq.	585	2713	577	2278	101	2278	1143
NT SEQ ID NO:	76	129	77	78	130	131	70
Vector	Uni-ZAP XR	pBluescript	Uni-ZAP XR				
ATCC Deposit No: Z and Date	97904 02/26/97 209050 05/15/97	97904 02/26/97 209050 05/15/97	97904 02/26/97 209050 05/15/97	97904 02/26/97 209050 05/15/97	97904 02/26/97 209050 05/15/97	97976 04/04/97	97904 02/26/97 209050 05/15/97
cDNA Clone ID	HTXCT40	HTXCT40	HRGDF73	HRDBF52	HRDBF52	HKMND45	HPEBD70
Gene No.	99	уу	29	××	89	89	69

Last AA of OR F	52	92
NT of AA First Last Predicted of S' NT First SEQ AA AA First AA First AA Last NO: NT Seq. Seq. Start Signal NO: Sig Sig Sig Scretcd of AA Seq. Codon Pep Y Pep Pep Portion OR F	28	36
Last AA of Sig Pep	27	25
First AA of Sig Pep	_	_
AA SEQ ID NO: Y	588 255 1	203
5' NT of First AA of Signal Pep		132 203
5' NT of Start Codon	588	557 132
3. NT of Clone Seq.	1043	557
S' NT of Clone Seq.	\$3\$	_
Total NT Seq.	1088	557
SEQ NÖ:	132	08
Vector	97904 Uni-ZAP XR 132 1088 535 1043 588 12/26/97 209050 15/15/97	97904 Uni-ZAP XR 80 22/26/97 209050 05/15/97
ATCC Deposit No: Z and Date	97904 02/26/97 209050 05/15/97	97904 02/26/97 209050 05/15/97
cDNA Clone ID	HPEBD70	70 HMCAB89
Gene No.	69	70

10

15

20

25

30

35

The Grant A

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences. reflected by the nucleotide position indicated as "5" NT of Clone Seq." and the "3" NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5" NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5" NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

10

15

20

25

30

35

1. 1. Garage 24 A.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below).

10

15

20

25

30

35

55,50000 9899348A)

It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

Signal Sequences

Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, supra.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., + or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely

10

15

20

25

The Age of the Season A

uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Polynucleotide and Polypeptide Variants

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

"Identity" per se has an art-recognized meaning and can be calculated using published techniques. (See, e.g.: (COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., ed., Oxford University Press, New York, (1988); BIOCOMPUTING: INFORMATICS AND GENOME PROJECTS, Smith, D.W., ed., Academic Press, New York, (1993); COMPUTER ANALYSIS OF SEQUENCE DATA, PART I,

- Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, (1994); SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heinje, G., Academic Press, (1987); and SEQUENCE ANALYSIS PRIMER, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, (1991).) While there exists a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans. (Carillo, H., and Lipton, D., SIAM J Applied Math 48:1073 (1988).) Methods commonly employed to determine identity or
- similarity between two sequences include, but are not limited to, those disclosed in "Guide to Huge Computers," Martin J. Bishop, ed., Academic Press, San Diego, (1994), and Carillo, H., and Lipton, D., SIAM J Applied Math 48:1073 (1988).
- Methods for aligning polynucleotides or polypeptides are codified in computer programs, including the GCG program package (Devereux, J., et al., Nucleic Acids Research (1984) 12(1):387 (1984)), BLASTP, BLASTN, FASTA (Atschul, S.F. et al., J. Molec. Biol. 215:403 (1990), Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park,
- 575 Science Drive, Madison, WI 53711 (using the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981).)

10

15

20

25

30

35

ENSIGNED PWC 9ENGLISHE

When using any of the sequence alignment programs to determine whether a particular sequence is, for instance, 95% identical to a reference sequence, the parameters are set so that the percentage of identity is calculated over the full length of the reference polynucleotide and that gaps in identity of up to 5% of the total number of nucleotides in the reference polynucleotide are allowed.

A preferred method for determing the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. 6:237-245) (1990).) The term "sequence" includes nucleotide and amino acid sequences. In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB search of a DNA sequence to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, and Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, and Window Size=500 or query sequence length in nucleotide bases, whichever is shorter. Preferred parameters employed to calculate percent identity and similarity of an amino acid alignment are: Matrix=PAM 150, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty=0.05, and Window Size=500 or query sequence length in amino acid residues, whichever is shorter.

As an illustration, a polynucleotide having a nucleotide sequence of at least 95% "identity" to a sequence contained in SEQ ID NO:X or the cDNA contained in the deposited clone, means that the polynucleotide is identical to a sequence contained in SEQ ID NO:X or the cDNA except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the total length (not just within a given 100 nucleotide stretch). In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to SEQ ID NO:X or the deposited clone, up to 5% of the nucleotides in the sequence contained in SEQ ID NO:X or the cDNA can be deleted. inserted, or substituted with other nucleotides. These changes may occur anywhere throughout the polynucleotide.

Further embodiments of the present invention include polynucleotides having at least 85% identity, more preferably at least 90% identity, and most preferably at least 95%, 96%, 97%, 98% or 99% identity to a sequence contained in SEQ ID NO:X or the cDNA contained in the deposited clone. Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the polynucleotides having at least 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity

10

15

20

25

30

35

GENGLIAF AT

will encode a polypeptide identical to an amino acid sequence contained in SEQ ID NO:Y or the expressed protein produced by the deposited clone.

Similarly, by a polypeptide having an amino acid sequence having at least, for example, 95% "identity" to a reference polypeptide, is intended that the amino acid sequence of the polypeptide is identical to the reference polypeptide except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the total length of the reference polypeptide. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

Further embodiments of the present invention include polypeptides having at least 80% identity, more preferably at least 85% identity, more preferably at least 90% identity, and most preferably at least 95%, 96%, 97%, 98% or 99% identity to an amino acid sequence contained in SEQ ID NO:Y or the expressed protein produced by the deposited clone. Preferably, the above polypeptides should exhibit at least one biological activity of the protein.

In a preferred embodiment, polypeptides of the present invention include polypeptides having at least 90% similarity, more preferably at least 95% similarity, and still more preferably at least 96%, 97%, 98%, or 99% similarity to an amino acid sequence contained in SEQ ID NO:Y or the expressed protein produced by the deposited clone.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as E. coli).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an

10

15

20

25

30

35

organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make

10

15

20

25

30

35

1 - 1

phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp. and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

10

15

20

25

30

35

ENCY TANK GENERAL

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967): Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

Polynucleotide and Polypeptide Fragments

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt. and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, and 701 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, and 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about"

includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

5

10

15

20

25

30

35

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983).)

10

15

20

25

30

35

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4.631.211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984): Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')2 fragments) which are capable of specifically binding to protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

10

15

20

25

30

35

Contract At

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In

preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the claimed invention.

Vectors, Host Cells, and Protein Production

5

10

15

20

25

30

35

and the second A

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli. Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS,

293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech. Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

5

10

15

20

25

30

35

GENERAL CENTRAL

1 . .

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

10

15

20

25

30

35

Assure of the second

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage

10

15

20

25

30

35

analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick. Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library).) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the

10

15

20

25

30

35

present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

10

15

20

25

30

35

11.00

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell . Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (1251, 1211), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 1311, 112In, 99mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20

10

15

20

25

30

35

millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing infiammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention could be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules

PCT/US98/04482

may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activity

5

10

15

20

25

30

35

A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome. lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from

10

15

20

25

30

35

Alberta Colors

inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitis, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease. Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

15

20

25

30

35

Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstron's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g.,

Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiollitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS).

pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps. Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium. Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceac.

Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter.

Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter.
Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis.
Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menigococcal), Pasteurellacea Infections (e.g., Actinobacillus,

30 Heamophilus, Pasteurella). Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease.

respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning.

Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria.

10

15

20

25

30

35

Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related). Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See. Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteocarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vascular (including vascular endothelium), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stoke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease. Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

20

25

30

5

10

15

Chemotaxis

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotaxic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, cosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotaxic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotaxic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

35

40.00

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat

disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

5

10

15

20

25

30

35

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Dresophila, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The

antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity, and (b) determining if a biological activity of the polypeptide has been altered.

15

10

5

Other Activities

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

20

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color. skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

25

30

35

4 وينان € ويا

. . .

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, caricadic rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities

qualities.

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

10

15

20

25

30

35

Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

10

15

20

25

30

35

ENGLY TO LOS

Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1. which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining

WO 98/39446 PCT/US98/04482

5

10

15

20

25

30

35

whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

10

15

20

25

30

35

Also preterred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the

15

20

25

30

35

amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an

10

15

20

25

30

35

Granite A

amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1: and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least

WO 98/39446

5

10

15

20

25

30

35

45 36 %

90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1: and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated

10

15

30

35

polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

Examples

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

	Vector Used to Construct Library	Corresponding Deposited Plasmid
	Lambda Zap	pBluescript (pBS)
20	Uni-Zap XR	pBluescript (pBS)
	Zap Express	pBK
	lafmid BA	plafmid BA
	pSport1	pSport1
	pCMVSport 2.0	pCMVSport 2.0
25	pCMVSport 3.0	pCMVSport 3.0
	pCR [®] 2.1	pCR [®] 2.1

Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988): Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-l Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3 primer sequences which flank the polylinker region ("S" is for Sacl and "K" is for KpnI which

WO 98/39446 PCT/US98/04482

119

are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the fl origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the fl ori generates sense strand DNA and in the other, antisense.

Vectors pSport 1, pCMVSport 2.0 and pCMVSport 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1 Blue. Vector pCR⁶2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

10

15

20

25

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids. each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with ³²P-γ-ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above.

15

20

25

30

35

BORT TO LA GRADULE

The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 µl of reaction mixture with 0.5 µg of the above cDNA template. A convenient reaction mixture is 1.5-5 mM MgCl₂, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source. although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then

10

15

20

25

30

35

be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

Example 3: Tissue Distribution of Polypeptide

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others. Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P³² using the rediprimeTM DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100TM column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHybTM hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

Example 4: Chromosomal Mapping of the Polynucleotides

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This

15

20

25

30

primer set is then used in a polymerase chain reaction under the following set of conditions: 30 seconds, 95°C: 1 minute, 56°C: 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

10 Example 5: Bacterial Expression of a Polypeptide

A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

The pQE-9 vector is digested with BamHl and Xbal and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacl repressor and also confers kanamycin resistance (Kan^r). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.⁶⁰⁰) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacl repressor, clearing the P/O leading to increased gene expression.

WO 98/39446 PCT/US98/04482

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

10

15

20

25

30

35

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl. 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM immidazole. Immidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number XXXXXX). This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence. 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (laclq). The origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with Ndel and XbaI, BamHI. XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for Ndel (5' primer) and XbaI, BamHI. XhoI, or

Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

5

10

15

20

25

30

1 . 4

Example 6: Purification of a Polypeptide from an Inclusion Body

The following alternative method can be used to purity a polypeptide expressed in $E\ coli$ when it is present in the form of inclusion bodies. Unless otherwise specified. all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* termentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris. 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfuidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive

10

15

Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGL.

Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A₂₈₀ monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 µg of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

20

25

30

35

Example 7: Cloning and Expression of a Polypeptide in a Baculovirus Expression System

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription,

15

20

25

30

35

Marine Congression

translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., Virology 170:31-39 (1989).

Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al.. "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures." Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five μg of a plasmid containing the polynucleotide is co-transfected with 1.0 μg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen. San Diego, CA), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One μg of BaculoGold™ virus DNA and 5 μg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 μl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 μl Lipofectin plus 90 μl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate

15

20

25

30

35

and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

To verify the expression of the polypeptide. Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 μ Ci of ³⁵S-methionine and 5 μ Ci ³⁵S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

Example 8: Expression of a Polypeptide in Mammalian Cells

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

15

20

25

30

35

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSport 2.0, and pCMVSport 3.0. Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1. Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem, et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991). Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No.209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, Xbal and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the

secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five µg of the expression plasmid pC6 is cotransfected with 0.5 µg of the plasmid pSVneo using lipofectin (Felgner et al., supra). The plasmid pSV2-neo contains a dominant selectable marker, the neo gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of metothrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 µM, 2 µM, 5 µM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 -200 µM. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Example 9: Protein Fusions

5

10

15

20

25

30

35

These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394.827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the

15

20

Christian Carry

activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHl cloning site. Note that the 3' BamHl site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHl, hnearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHl site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

Human lgG Fc region:

GGGATCCGGAGCCCAAATCTTCTGACAAAACTCACACATGCCCACCGTGCC CAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCCAAAACC 25 CAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTCACATGCGTGGTGGT GGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACG GCG1GGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAAC AGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTG AATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAACCCCC 30 ATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGT GTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCT GACCTGCCTGGTCAAGGCTTCTATCCAAGCGACATCGCCGTGGAGTGGGA GAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGG 35 ACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCA GGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGC

10

15

20

25

30

35

ACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAGTGCGACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

Example 10: Production of an Antibody from a Polypeptide

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 μg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a

mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F(ab')2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4.816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

Example 11: Production Of Secreted Protein For High-Throughput Screening Assays

25

30

35

عاملت المراجع المراجع

10

15

20

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at 2 x 10^s cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

5

10

15

20

25

30

35

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem l complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (see below) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other

proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

5 HGS-CHO-5 medium formulation:

Inorganic Salts

CaCl2 (anhyd)	116.6 mg/L
CuSO ₄ -5H ₂ O	0.00130
Fe(NO ₃) ₃ -9H ₂ O	0.050
FeSO ₄ -7H ₂ O	0.417
KCl	311.80
MgCl.	28.64
MgSO ₄	48.84
NaCl	6995.50
NaHCO _z	2400.0
NaH ₂ PO ₄ -H ₂ 0	62.50
Na ₂ HPO4	71.02
ZnSO ₄ -7H ₂ O	.4320

Lipids

Arachidonic Acid	.002 mg/L
Cholesterol	1.022
DL-alpha-	.070
Tocopherol-Acetate	
Linoleic Acid	0.0520
Linolenic Acid	0.010
Myristic Acid	0.010
Oleic Acid	0.010
Palmitric Acid	0.010
Palmitic Acid	0.010
Pluronic F-68	100
Stearic Acid	0.010
Tween 80	2.20

10 Carbon Source

D-Glucose	4551 mg/L

Amino Acids

L- Alanine	130.85 mg/ml
L-Arginine-HCL	147.50
L-Asparagine-H ₂ 0	7.50

L-Aspartic Acid	6.65
L-Cystine-2HCL-	29.56
H ₂ 0	
L-Cystine-2HCL	31.29
L-Glutamic Acid	7.35
L-Glutamin€	365.0
Glycine	18.75
L-Histidine-HCL-	52.48
H ₂ 0	
L-Isoleucine	106.97
L-Leucine	111.45
L-Lysine HCL	163.75
L-Methionin€	32.34
L-Phenylalainine	68.48
L-Proline	40.0
L-Serine	26.25
L-Threonine	101.05
L-Tryptophan	19.22
L-Tryrosine-2Na-	91.79
2H ₂ 0	
L-Valine	99.65

Vitamins

Biotin	0.0035 mg/L
D-Ca Pantothenate	3.24
Choline Chlorid€	11.78
Folic Acid	4.65
i-Inositol	15.60
Niacinamide	3.02
Pyridoxal HCL	3.00
Pyridoxine HCL	0.031
Riboflavin	0.319
Thiamine HCL	3.17
Thymidine	0.365
Vitamin B ₁	0.680

Other Components

HEPES Buffer	25 mM
Na Hypoxanthine	2.39 mg/L
Lipoic Acid	0.105
Sodium Putrescine-2HCL	0.081
Sodium Pyruvate	55.0
Sodium Selenite	0.0067
Ethanolamine	20uM
Ferric Citrate	0.122
Methyl-B-Cyclodextrin complexed with	41.70
Linoleic Acid	

10

15

20

25

30

Methyl-B-Cyclodextrin complexed with Oleic Acid	33.33
Methyl-B-Cyclodextrin complexed with Retinal Acetate	10

Adjust osmolarity to 327 mOsm

Example 12: Construction of GAS Reporter Construct

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2. Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo. PRL. GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proxial region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

WO 98/39446 PCT/US98/04482

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

5

11.07 () Vi () OF FALLY A

	Y : 1	4	JAKs Jaka	11.0	Lale?	<u>STATS</u>	GAS(elements) or ISRF.
	Ligand	tyk2	<u>Jak 1</u>	<u>Jak2</u>	<u>Jak3</u>		
5	IFN family IFN-a/B IFN-ξ II-10	+ +	+ + ?	- + ?	-	1,2,3 1 1,3	ISRE GAS (IRF1>Lys6>IFP)
10	gp130 family IL-6 (Pleiotrohic) Il-11(Pleiotrohic) OnM(Pleiotrohic)	+ ? ?	+ + +	+ ? +	? ? ?	1.3 1.3 1.3	GAS (IRF1>Lys6>IFP)
15	LIF(Pleiotrohic) CNTF(Pleiotrohic) G-CSF(Pleiotrohic) IL-12(Pleiotrohic)	? -/+ ? +	+ + + -	+ + ? +	???	1.3 1.3 1.3 1.3	
20	g-C family IL-2 (lymphocytes) IL-4 (lymph/mycloid) IL-7 (lymphocytes) IL-9 (lymphocytes) IL-13 (lymphocyte) IL-15	- - - - ?	+ + + + +	- ?	+ + + + + + ? + +	1.3.5 6 5 5 6 5	GAS GAS (IRF1 = IFP >>Ly6)(IgH) GAS GAS GAS GAS GAS
25	gp140 family IL-3 (myeloid) IL-5 (myeloid) GM-CSF (myeloid)	-	- -	+ + +	-	5 5 5	GAS (IRF1>IFP>>Ly6) GAS GAS
30		• •					
35	Growth hormone fam GH PRI EPO	ily ? ? ?	- +/- -	+ + + +	-	5 1,3.5 5	GAS(B-CAS>IRF1=IFP>>Ly6)
J.	Receptor Tyrosine Ki EGF PDGF CSF-1	nases ? ?	+ 	+ +	- -	1.3 1.3 1.3	GAS (IRF1) GAS (not IRF1)
40	.	•					

10

15

20

25

30

35

1 11

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an Xhol site. The sequence of the 5' primer is: 5':GCGCCTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCCGAAATGATTTCCCCCGAAATGATTTCCCCCGAAATGATTTCCCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

5':CTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCGAAATG
ATTTCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCC
CTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGC
CCCATGGCTGACTAATTTTTTTTATTTATTTATGCAGAGGCCGAGGCCGCCTCGGC
CTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTT
TGCAAAAAAGCTT:3' (SEQ ID NO:5)

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

15

20

25

3()

35

PACE TOWN WITHOUT A

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using Sall and Notl, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, Il-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell). Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

Example 13: High-Throughput Screening Assay for T-cell Activity.

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliterate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml genticin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1% Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies)

10

15

20

25

30

with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells (10° per transfection), and resuspend in OPTI-MEM to a final concentration of 10° cells/ml. Then add 1ml of 1 x 10° cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPM1 + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Genticin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100, 000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

15

20

25

30

Example 14: High-Throughput Screening Assay Identifying Myeloid Activity

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest 2x10e⁷ U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM Na₂HPO₄.7H₂O, 1 mM MgCl₂, and 675 uM CaCl₂. Incubate at 37°C for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

These cells are tested by harvesting $1x10^8$ cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of $5x10^5$ cells/ml. Plate 200 ul cells per well in the 96-well plate (or $1x10^5$ cells/well).

Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

Example 15: High-Throughput Screening Assay Identifying Neuronal 35 Activity.

10

15

20

25

30

35

GE WALLER L

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes. EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat phenochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor). The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)

5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes Xhol/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heatinactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS (Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as $5x10^5$ cells/ml.

Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to $1x10^5$ cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

15

20

25

30

10

5

Example 16: High-Throughput Screening Assay for T-cell Activity

NF-κB (Nuclear Factor κB) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF-κB regulates the expression of genes involved in immune cell activation, control of apoptosis (NF-κB appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF- κB is retained in the cytoplasm with I- κB (Inhibitor κB). However, upon stimulation, I- κB is phosphorylated and degraded. causing NF- κB to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF- κB include IL-2, IL-6, GM-CSF, ICAM-1 and class 1 MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF-kB promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF-kB would be useful in treating

10

15

20

diseases. For example, inhibitors of NF-kB could be used to treat those diseases related to the acute or chronic activation of NF-kB, such as rheumatoid arthritis.

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:

5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene)

Sequencing with the T7 and T3 primers confirms the insert contains the following

Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

5':CTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCC ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCA TCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACT AATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTC CAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTT: 3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2-promoter plasmid (Clontech) with this NF-κB/SV40 fragment using XhoI and HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF-κB/SV40/SEAP

cassette is removed from the above NF-κB/SEAP vector using restriction enzymes Sall and NotI, and inserted into a vector containing neomycin resistance. Particularly, the NF-κB/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

15

20

25

ENCOUNT OF A VICTOR OF A MADE

Once NF-κB/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described in Example 13. As a positive control, exogenous TNF alpha (0.1,1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17: Assay for SEAP Activity

As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution. Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 µl of 2.5x dilution buffer into Optiplates containing 35 µl of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 µl Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 µl Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

Reaction Ruffer Formulation:

Reaction B	uller Formulation:	
# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
] 4	80	4
15	85	4.25
16	9()	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6

PCT/US98/04482

5

10

15

23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25
28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	Ò
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10.
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-3, used here.

For adherent cells, seed the cells at 10.000 -20.000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO₂ incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-3 is made in 10% pluronic acid DMSO. To load the cells with fluo-3, 50 ul of 12 ug/ml fluo-3 is added to each well. The plate is incubated at 37°C in a CO₂ incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10⁶ cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-3 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10⁶ cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-3. The supernatant is added to the well, and a change in fluorescence is detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling even which has resulted in an increase in the intracellular Ca⁺⁺ concentration.

20

25

30

35

Ethic () which decrease A)

5

10

15

Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

10

15

20

25

30

35

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately 25.000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford,MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford,MA) are used to cover the Loprodyne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyne plates (20,000/200ml/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na3VO4, 2 mM Na4P2O7 and a cocktail of protease inhibitors (# 1836170) obtained from Boeheringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16.000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a

10

15

20

25

15511 1 . .

الروان والماعل

biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg₂₊ (5mM ATP/50mM MgCl₂), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl₂, 5 mM MnCl₂, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the components gently and preincubate the reaction mix at 30°C for 2 min. Initial the reaction by adding 10ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mm EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavadin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phospotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as above.

Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity

As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase.

Src. Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other

phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1 and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4° C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

25

30

35

1 - 1 - 1 - 1 - 1 - 1 - 1

10

15

20

Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products is then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies).

10

15

20

25

30

35

ENCY TORK WHERE

The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals is identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenindeoxy-uridine 5'-triphosphate (Boehringer Manheim), and FISH performed as described in Johnson. Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology. Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect soluble polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method

10

15

20

25

30

35

. . . .

described in Example 10. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

Example 23: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 µg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 µg/kg/hour to about 50 µg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracistemally, intravaginally,

15

20

25

30

35

intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), bucally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or mirocapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2- hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52.322: EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

WO 98/39446 PCT/US98/04482

155

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

And the second of the

5

10

15

20

25

10

15

20

25

30

1. . .

Example 24: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

Example 25: Method of Treating Increased Levels of the Polypeptide

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5. 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

Example 26: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin, is added. The flasks are then incubated at 37°C for approximately one week.

WO 98/39446 PCT/US98/04482

157

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is being produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

THE CHI PWC GENERAL

10

15

20

25

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference.

	(1) GENERAL INFORMATION:
	(i) APPLICANTS: Human Genome Sciences, Inc. et al.
	(ii) TITLE OF INVENTION: 70 Human Secreted Proteins
5	(iii) NUMBER OF SEQUENCES: 273
	(iv) CORRESPONDENCE ADDRESS:
	(A) ADDRESSEE: Human Genome Sciences, Inc.
	(B) STREET: 9410 Key West Avenue
	(C) CITY: Rockvill€
10	(D) STATE: Maryland
	(E) COUNTRY: USA
	(F) ZIP: 20850
	(v) COMPUTER READAELE FORM:
15	(A) MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
	(B) COMPUTER: HP Vectra 486/33
	(C) OPERATING SYSTEM: MSDOS version 6.2
	(D) SOFTWARE: ASCII Text
	(vi) CURRENT APPLICATION DATA:
20	(A) APPLICATION NUMBER:
	(B) FILING DATE: March 6, 1998
	(C) CLASSIFICATION:
	(vii) PRIOR APPLICATION DATA:
	(A) APPLICATION NUMBER:
25	(B) FILING DATE:
	(viii) ATTORNEY/AGENT INFORMATION:
	(A) NAME: A. Anders Brookes
	(B) REGISTRATION NUMBER: 36,373

(C) REFERENCE/DOCKET NUMBER: PS001PCT

20

25

(vi)	TELECOMMUNICATION	INFCHMATION:

(A) TELEPHONE: (301) 309-8504

(E) TELEFAX: (301) 309-8439

5 (2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 733 base pairs

(E) TYPE: nucleic acid

(C) STRANDEDNESS: double

10 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGGATCCGGA GCCCAAATCT TCTGACAAAA CTCACACATG CCCACCGTGC CCAGCACCTG 60 AATTOGAGGG TGUACOGTOA GTOTTCCTOT TCCCCCCAAA ACCCAAGGAC ACCCTCATGA 120 180 TOTOCOGGAC TOCTMAGETC ACATGCGTGG TGGTGGACGT AAGCCACGAA GACCCTGAGG TCAAGTTCAA CTGGTACGTG GACGGCGTGG AGGTGCATAA TGCCAAGACA AAGCCGCGGG 240 AGGAGCAGTA CAACAGCACG TACCGTGTGG TCAGCGTCCT CACCGTCCTG CACCAGGACT 300 360 GGCTGAATGG CAAGGAGTAC AAGTGCAAGG TCTCCAACAA AGCCCTCCCA ACCCCCATCG AGAAAACCAT CTCCAAAGCC AAAGGGCAGC CCCGAGAACC ACAGGTGTAC ACCCTGCCCC 420 CATCCCGGGA TGAGCTGACC AAGAACCAGG TCAGCCTGAC CTGCCTGGTC AAAGGCTTCT 480 ATCCAAGCGA CATCGCCGTG GAGTGGGAGA GJAATGGGCA GCCGGAGAAC AACTACAAGA 540 CCACGCCTCC CGTGCTGGAC TCCGACGGCT CCTTCTTCCT CTACAGCAAG CTCACCGTGG 600 660 ACAAGAGCAG GT990AG0AG GGGAACGTCT TOTCATGCTO CGTGATGCAT GAGGCTCTGC 720 ACAACCACTA CACGCAGAAG AGCCTCTCCC TSTCTCCGGG TAAATGAGTG CGACGGCCGC 733 GACTCTAGAG GAT

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids

30 (B) TYPE: amino acid

Optional A

	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:	
	Trp Ser Xaa Trp Ser	
5	1 5	
	(2) INFORMATION FOR SEQ ID NO: 3:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 86 base pairs	
10	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:	
15	GCGCCTCGAG ATTTCCCCGA AATCTAGATT TCCCCGAAAT GATTTCCCCG AAATGATTTC	60
	CCCGAAATAT CTGCCATCTC AATTAG	86
	(2) INFORMATION FOR SEQ ID NO: 4:	
	(i) SEQUENCE CHARACTERISTICS:	
20	(A) LENGTH: 27 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:	
25	GCGGCAAGCT TTTTGCAAAG CCTAGGC	27
	(2) INFORMATION FOR SEQ ID NO: 5:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 271 base pairs	
30	(B) TYPE: nucleic acid	

BNC CONTRACT SERVICERAL

	(C) STRANDEDNESS: GCUDIE	
	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:	
	CTCGAGATTI CCCCGAAATC TAGATTTCCC CGAAATGATT TCCCCGAAAT GATTTCCCCG	60
5	AAATATCTGC CATCTCAATT AGTCAGCAAC CATAGTCCCG CCCCTAACTC CGCCCATCCC	120
	GCCCCTAACT CCGCCCAGTT CCGCCCATTC TCCGCCCCAT GGCTGACTAA TTTTTTTAT	180
	TTATGCAGAG GCCGAGGCCG CCTCGGCCTC TGAGCTATTC CAGAAGTAGT GAGGAGGCTT	240
	TTTTGGAGGC CTAGGCTTTT GCAAAAAGCT T	271
10	(2) INFORMATION FOR SEQ ID NO: ϵ :	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 32 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: doubl€	
15	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:	
	GCGCTCGAGG GATGACAGCG ATAGAACCCC GG	32
20		
20	(2) INFORMATION FOR SEQ ID NO: 7:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 31 base pairs	
	(B) TYPE: nucleic acid	
25	(C) STRANDEDNESS: doubl∈ (D) TOPOLOGY: linear	
23		
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:	
	GCGAAGCTTC GCGACTCCCC GGATCCGCCT C	31
	ccention occurrence oursecourt	2.1
30	(2) INFORMATION FOR SEQ ID NO: E:	

	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 12 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
5	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:	
	GGGGACTTTC CC	12
10	(2) INFORMATION FOR SEQ ID NO: 9:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 73 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
15	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:	
	GCGGCCTCGA GGGGACTTTC CCGGGGACTT TCCGGGGACT TTCCATCCTG	60
	CCATCTCAAT TAG	73
20		
	(a) ANTIONNETON FOR CEO TO NO. 16.	
	(2) INFORMATION FOR SEQ ID NO: 10: (i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 256 base pairs	
25	(B) TYPE: nucleic acid	
4 J	(C) STRANDEDNESS: doubl∈	
	(D) TOPOLOGY: linear	
	(5) 30,02001	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:	
30	CTCGAGGGA CTTTCCCGGG GACTTTCCG GGACTTTCCA TCTGCCATCI	60

CTTTTGCAAA	AAGCTI					25€
GGCCGCCTCG	GCCTCTGAGC	TATTCCAGAA	GTAGTGAGGA	GGCTTTTTTS	GAGGCCTAGG	240
CAGTTCCCC	CATTCTCCGC	CCCATGGCTG	ACTAATTTT	ETATTTATUT	CAGAGGCCGA	180
CAATTAGTCA	GCAACCATAG	DOCCOSOCCI	AACTCCGCCC	ETCCCCCCCCC	TAACTCCGCC	120

5

10

15

20

25

30

(2) INFORMATION FOR SEC ID NO: 11:

(i) SEQUENCE CHARACTEFISTICS:

(A) LENGTH: 1739 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GEGETCENGA GGEOGEGGG CETYCEAGAGA GGACAGCOG CETYCENGG GACATGCGGC 60 CCCAGGAGCT CCCCAGGCTC GCGTTCCCGT TGCTGCTGTT GCTGTTGCTG CTGCTGCCGC CURCOGUERTO COCTOCCCAC AGUACULACOC GTTVCGACCO CACCUGURGAS TOUCTGGACG 180 240 CCCGCCAGCT GCCCGCGTGG TTTGACCAGG CCAAGTTCGG CATCTTCATC CACTGGGGAG TGTTTTCCGT GCCCAGCTTC GGTAGGGAGT GGTTCTGGTG GTATTGGCAA AAGGAAAAGA 300 TACCGAAGTA TGTGGAATTT ATGAAAGATA ATTACCCTCC TARTTCAAA TATGAAGATT 360 TTGGACCACT ATTTACAGCA AAATTTTTTA ATGCCAACCA RTGGGCARAT ATTTTYCAGG CCTCTGGTGC CAAATACATT GTCTTAACTT CCAAACATCA TGAAGGCTTT ACCTTGTGGG 480 GGTCAGAATA TTCGTGGAAC TGGAATGCCA TAGATGAGGG GCCCAAGAGG GACATTGTCA 540 AGGAACTTGA GGTAGCCATT AGGAACAGAA CTGACCTGCG TTTTGGACTG TACTATTCCC 600 TTTTTSAATG GTTTCATCCG CTCTTCCTTG AGBATGAATC CAGTTCATTC CATAAGCGGC 660 72: AATTTCCAGT TTCTAAGACA TTGCCAGAGC TCTATGAGTT AGTGAACAAC TATCAGCCTG AGGTTOTGTG GTCGGATGGT GACGGAGAG CACCGGATCA ATACTGGAAC ANGACAGGCI 780 TOTTOGCCTG GITATATAAT GAAAGOCCAG TTCGGGGGCAC AGTAGICAGO AATGATCGTT GGGGAGCTGG TAGCATCTGT AAGCATGGTG GCTTCTATAC CTGCAGTGAT CGTTATAACC 900 CAGGACATCT TTTGCCACAT AAAT933AAA ACTGCATGAC AATAGACAAA CTGTCCTGGG 960 GCTATAGGAG GGAAGCTGGA ATCTCTGACT ATCTTACAAT TGAAGAATTG GTGAAGCAAC 1020

PCT/US98/04482

Particle of Land

	TIGTAGAGAC	AGTTTCATGT	GGAGGAAATC	TTTTGATGAA	TATTGGGCCC	ACACTAGATG	1080
	GCACCATTTC	TGTAGTTTTT	GAGGAGCGAC	TGAGGCAAAT	GGGGTCCTGG	CTAAAAGTCA	1140
	ATGGAGAAGC	TATTTATGAA	ACCCATACCT	GGCGATCCCA	GAATGACACT	GTCACCCCAG	1200
	ATGTGTGGTA	CACATCCAAG	CCTAAAGAAA	AATTAGTCTA	TGCCATTTTT	CTTAAATGGC	1260
5	CCACATCAGG	ACAGCTGTTC	CTT3GCCATC	CCAAAGCTAT	TCTGGGGCA	ACAGAGGTGA	1320
	AACTACTGGG	CCATGGACAG	CCACTTAACT	GGATTTCTTT	GGAGCAAAAT	GGCATTATGG	1380
	TAGAACTGCC	ACAGCTAACC	ATTCATCAGA	TGCCGTGTAA	ATGGGGCTGG	GCTCTAGCCC	1440
	TRACTAATGT	GATCTAAAGT	GCAGCAGAGT	GGCTGATGCT	GCAAGTTATG	TCTAAGGCTA	1500
	GGAACTATCA	GGTGTCTATA	ATTGTAGCAC	ATGGAGAAAG	CAAATGTAAA	ACTGGATAAG	1560
10	TTTATTAAAA	TGGCAGTTCA	GCCCTTTCCC	TTTTTCCCAC	TTTTTTAAAT	CTTAAATTAC	1620
	CCATGTAACC	ATTTTAACTC	TCCAGTGCAC	TTTGCCATTA	AAGTCTCTTC	ACATTGAAAA	1680
	АААААААА	AAAAACCCCG	CGGGGGGGC	CCGGGNACCC	CATTTCGCCC	NTAAAGGGG	1739

(2) INFORMATION FOR SEC ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 844 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12: 20 GGCCCCTGGG CCCGAGGGGC TGGAGCCGGG CCGGGGGGAT GTGGAGCGCG GGCCGCGGGG 60 GGGCTGCCTG GCCGGTGCTG TTGGGGGCTGC TGCTGGCGGT GTTAGTGCCG GGCGGTGGTG 120 CCGCCAAGAC CGGTGCGGAG CTCGTGACCT GCGGGTCGGT GCTGAAGCTG CTCAATACGC 180 ACCACCGCGT GCGGCTGCAC TCGCACGACA TCAAATACGG ATCCGGCAGC GGCCAGCAAT 240 CGGTGACCGG CGTAGAGGCG TCGGACGACG CCAATAGCTA CTGGCGGATC CGCGGCGGCT 300 25 CGBAGGGGG GTGCCGCCGC GGGTCCCCGG TGCGCTGCGG GCAGGCGGTG AGGCTCACGC 360 ATTETIGETTAC GEGCAAGAAC CTGCACACGC ACCACTTCCC GTCGCCGETG TCCAACAACC 420 AGGAGGTGAG TGCCTTTGGG GAAGACGGCG AGGGCGACGA CCTGGACCTA TGGACAGTGC 480 GCTGCTCTGG ACAGCACTGG GAGCGTGAGG CTGCTGTGCG CTTCCAGCAT GTGGGCACCT 540 600 30 CTGTGTTCCT GTCAGTCACG GGTGAGCAGT ATGGAAGCCC CATCCGTGGG CAGCATGAGG

	TOCACGGCAT GOOGAGTGOG AACACGCACA ATACGTGGAA GGCATCATCA	660
	TCAAGCCTAG TETEGRAGOOD TOTEGRAGGTC ACGATGAACT OTWAGTETGT GGATGGATGG	720
	GTGGATGGAG GSTGGCAGST GSGSGCSTCTG CAGGGCCACT CTTYGGCAGAG ACTTYGGGTT	780
	AAAAAAAA AAAETATOTO OTTOTAADAA ATTAETOTYT OCOTOAAGTO CTODOODATOT	84(
5	AAAA	844
	(2) INFORMATION FOR SEQ ID NO: 13:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 776 base pairs	
10	(F) TYPE: nucleic acid	
	(C) STRANDEDNESS: doubl∈	
	(I) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:	
	TTCGAAATAA AAGATCT9CT CAAGAGAGCC GCAGAAAAAG AAGGTGTATG TTGGGGGTTT	6(
15	AGAGAGCAGG GTCTTGAAAT ACACAGCCCA GAATATGGAG CTTCAGAACA AAGTACAGCT	120
	TCTGGAGGAA CABAATTTBT CCCTTCTAGA TCAACTGAGG AAACTCCAGG CCATGBTGAT	180
	TGAGATATCA AACAAAACCA GCAGCAGCAG CACCTGCATC TN93TCCTAC TAGTCTCCTT	240
	CTGCCTCCTC CTRSTACTG CTATGTACTC CTCTGACACA AGGGGGAGCC TGCCAGCTGA	300
	GCATGSAGTG TTSTCCOSCC AGCTTCGTGC CCTCCCCAGI GAGSACCCTT ACCAGCTGGA	360
20	GCTGCCTGCC CT3CAGTCAG AAGT3CCGAA AGACAGCACA CACCAGTGGT TGGACGGCTC	42
	AGACTETETA CTOCAGECCO CTESCACAC TTCCTGCCTG CTECATTACA TGCCTCAGC	48
	TCCCAGTGCA GAGCCTCCCC TGGAGTGGCC ATTCCCTGAC CTCTTCTCAG AGCCTCTCTG	541
	COGAGETCCC ATTOTTOTTC THETAGGCAAA TOTCACAAGG AAGEGAGGAT GGCTTCCTAC	600
25	TOSTAGCCCC TOTSTOATTT TOOMOGACAG ATACTCAGGC TAGATATGAG GATATGTGGG	661
	GGGTCTCAGC AGGAGTCTGG GGGGGCAC ATCTGTGTCC AAATAAAAAG CGGTGGGCAA	72
	GGGCTGGCCG CAGUTTCTGT GCCUTGTCAG GACGACTGAG GGCTCAAACA CACCAC	776

(2) INFORMATION FOR SEQ ID NO: 14:

30 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1376 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:	
	GAATTCGGCA CGAGGCGCCT ACCCTGCCTG CAGGTGAGCA GTGGTGTGTG AGAGCCAGGC	6.C
	GTCCCTCTGC CTGCCCACTC AGTGGCAACA CCCGGGAGCT GTTTTGTCCT TTGTGGAGCC	120
	TCAGCAGTTC CCTCTTTCAG AACTCACTGC CAAGAGCCCT GAACAGGAGC CACCATGCAG	180
	TGCTTCAGCT TCATTAAGAC CATGATGATC CTCTTCAATT TGCTCATCTT TCTGTGTGGT	240
10	GCAGCCCTGT TGGCAGTGGG CATCTGGGTG TCAATCGATG GGGCATCCTT TCTGAAGATC	300
	TTCGGGCCAC TGTCGTCCAG TGCCATGCAG TTTGTCAACG TGGGCTACTT CCTCATCGCA	360
	GCCGGCGTTG TGGTCTTTGC TCTTGGTTTC CTGGGCTGCT ATGGTGCTAA GACTGAGAGC	420
	AAGTGTGCCC TCGTGACGTT CTTCTTCATC CTCCTCCTCA TCTTCATTGC TGAGGTTGCA	480
	GCTGCTGTGG TCGCCTTGGT GTACACCACA ATGGCTGAGC ACTTCCTGAC GTTGCTGGTA	540
15	GTGCCTGCCA TCAAGAAAGA TTATGGTTCC CAGGAAGACT TCACTCAAGT GTGGAACACC	60C
	ACCATGAAAG GGCTCAAGTG CTGTGGCTTC ACCAACTATA CGGATTTTGA GGACTCACCC	660
	TACTTCAAAG AGAACAGTGC CTTTCCCCCA TTCTGTTGCA ATGACAACGT CACCAACACA	720
	GCCAATGAAA CCTGCACCAA GCAAAAGGCT CACGACCAAA AAGTAGAGGG TTGCTTCAAT	780
	CAGCITTIGI ATGACATCCG AACTAATGCA GTCACCGTGG GTGGTGTGGC AGCTGGAATT	840
20	GGGGGCCTCG AGCTGGCTGC CATGATTGTG TCCATGTATC TGTACTGCAA TCTACAATAA	900
	GTCCACTTCT GCCTCTGCCA CTACTGCTGC CACATGGGAA CTGTGAAGAG GCACCCTGGC	960
	AAGCAGCAGT GATTGGGGGA GGGGACAGGA TCTAACAATG TCACTTGGGC CAGAATGGAC	1020
	CTGCCCTTTC TGCTCCAGAC TTGGGGCTAG ATAGGGACCA CTCCTTTTAN GCGATGCCTG	1080
	ACTITECTTC CATTGGTGGG TGGATGGGTG GGGGGCATTC CAGAGCCTCT AAGGTAGCCA	11 4 0
25	GTTCTGTTGC CCATTCCCCC AGTCTATTAA ACCCTTGATA TGCCCCCTAG GCCTAGTGGT	1200
	GATCCCAGTG CTCTACTGGG GGATGAGAGA AAGGCATTTT ATAGCCTGGG CATAAGTGAA	1260
	ATCAGCAGAG CCTCTGGGTG GATGTGTAGA AGGCACTTCA AAATGCATAA ACCTGTTACA	1320
	ATGTTRAAAA AAAAAAAAA AAAAAAAAA AAAAAAYTCG AGGGGGGTCC CGTACC	137€

	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 502 base pairs	
	(B) TYPE: nucleic ació	
	(C) STRANDEDNESS: doubl€	
5	(E) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:	
	TARARCASTS CONSCIONA AGEGAGACT CACTORATAT CTGTTGAATG AATGAATGAA	60
	TAATGOOTIS GETSAAGGAA TGAATGOOTIS AATGAATGAT TTOTOTITTO COTOGOCACI	120
	GTCTGSAGTO COCAGSACAG GCATGGGCAG CAGTOGCTGG TOTGTGGCCT GTCCCACTG	180
10	ACTIGODATO DECENDADA ADDIACITADA ESCENDETO ETENDE TACHE UNESCONDE	240
	TOTOCTICACY: GRANAROTHOU TOCOROCONA GOTOCOAGAA CICACTGOAS GETERREGRA	300
	AFARCAGARA CHATCTOCIGA GOCCOTGAGA DAGCOCAGAGA AGCCCAGGAGAGA	360
	AAATGCAGGT GTTGAGAGGA GTTTTGCCTT CTTTTTGAG TTGAATATGA GATTTCCGAG	420
	CAGCCATGAC GAGTTGGGTT GGTGGAGGGG GTGGAGGGTCA GATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	48(
15	GGGGTCCCCT TYGATCTCCT CT	502
	(2) INFORMATION FOR SEQ ID NO: 16:	
	(i) SEQUENCE CHARACTERISTICS	
	(A) LENGTH: 425 base pairs	
20	(E) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEC ID NO: 16:	
	ATCTCTATO ADTICCETTO TAREBULADA ADACOTOCOT DOCETICA CLEATOTITE	60
25	CTEGETT TETTEGERGE AGREET TETTEGERGE AGREET TETTEGERGE TETTEGERGE	120
	TGTGBATTTT TGCATCAGTG GGAAAACAAG AGGACAGAAG CCAAAACTTTG TGATTATTTT	180
	GECCGATGAC ATEEGGTEGG GEEGGCACTGG GEGGAAACAA AGGACACTGC	240
	CAACCTTGAT AAGATGGCTT CGGAGGGAAT GARGTGARTC TTGARATGCC ARGCCAGCTT	300
	TOTTTGBAWS TOTTACTOCC GTTOTTBAAA AGGGAAAGGG GOGTGCAAAG CACTTAARGA	360
30	WYCATKGATG GACCCATGTG ATTTARTTAA TTTATTAATT AATTTGGTTT GGAARCCAGC	420

WO 98/39446 PCT/US98/04482

169

ATAGC 425

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

5 (A) LENGTH: 1316 base pairs

The first part of the second of

(B) TYPE: nucleic acid

(C) STRANDEDNESS: doubl€

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

10	GGCACGAGGA	GCTGGGGGAG	CCTGAGGTGC	GCTACGTGGC	TGGCATGCAT	GGGAACGAGG	60
	CCCTGGGGCG	GGAGTTGCTT	CTGCTCCTGA	TGCAGTTCCT	GTGCCATGAG	TTCCTGCGAG	120
	GGAACCCACG	GGTGACCCGG	CTGCTCTCTG	AGATGCGCAT	TCACCTGCTG	CCCTCCATGA	180
	ACCCTGATGG	CTATGAGATC	GCCTACCACC	GGGGTTCAGA	GCTGGTGGGC	TGGGCCGAGG	240
	GCCGCTGGAA	CAACCAGAGC	ATCGATCTTA	ACCATAATTT	TGCTGACCTC	AACACACCAC	300
15	TGTGGGAAGC	ACAGGACGAT	GGGAAGGTGC	CCCACATCGT	CCCCAACCAT	CACCTGCCAT	360
	TGCCCACTTA	CTACACCCTG	CCCAATGCCA	CCGTGGCTCC	TGAAACGCGG	GCAGTAATCA	420
	AGTGGATGAA	GCGGATCCCC	TTTGTGCTAA	GTGCCAACCT	CCACGGGGGT	GAGCTCGTGG	480
	TGTCCTACCC	ATTCGACATG	ACTCGCACCC	CGTGGGCTGC	CCGCGAGCTC	ACGCCCACAC	540
	CAGATGATGC	TGTGTTTCGC	TGGCTCAGCA	CTGTCTATGC	TGGCAGTAAT	CTGGCCATGC	600
20	AGGACACCAG	CCGCCGACCC	TGCCACAGCC	AGGACTTCTC	CGTGCACGGC	AACATCATCA	660
	ACGGGGCTGA	CTGGCACACG	GTCCCCGGGA	GCATGAATGA	CTTCAGCTAC	CTACACACCA	720
	ACTGCTTTGA	GGTCACTGTG	GAGCTGTCCT	GTGACAAGTT	CCCTCACGAG	AATGAATTGC	780
	CCCAGGAGTG	GGAGAACAAC	AAAGACGCCC	TCCTCACCTA	CCTGGAGCAG	GTGCGCATGG	840
	GCATTGCAGG	AGTGGTGAGG	GACAAGGACA	CGGAGCTTGG	GATTGCTGAC	GCTGTCATTG	900
25	CCGTGGATGG	GATTAACCAT	GACGTGACCA	. CGGCCTGGGG	CGGGGATTAT	TGGCGTCTGC	960
	TGACCCCAGG	GGACTACATG	GTGACTGCCA	GTGCCGAGGG	CTACCATTCA	GTGACACGGA	1020
						ACCAAGACTC	
						GACCTTCGCA	
						AGCCCTAGGG	
20							126
30	CAGGUTGGAC	CIGTCAAGAC	, GGGAAGGAGA	ADADA L'UADA	1 CHURCHIA	AAGTGAGGAA	120

	AAGGTGCTCA TTAAAGCTAC CGGGCACCTT RAARAAAAA AAAAAAAAA AAAAAA	131€
	(2) INFORMATION FOR SEQ ID NO: 18:	
	(i) SEQUENCE CHAFACTERISTICS:	
5	(A) LENGTH: 436 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:	
10	OTRATATAER TATRAARACO TYTATRARAR AATAAAARTA TYATARRITAA OYTAAAAAA	60
	CACGTGTCAA GAAAAATTCA GTACACGAGC AAGAAGCCAT TAACTCTGAC CCAGAGTTGT	120
	CTAATTGTGA AAATTTTCAG AAGACTGATG TGAAAGATGA TCTGTCTGAT CCTCCTGTTG	180
	CAAGCAGTTG TATTTCTGAG AAGTCTCCAC GTAGTCCACA ACTTTCAGAT TMTGGACTTG	240
	AGCGGTACAT CGTATCCCAA GTTCTACCAA ACCCTCCACA GGCAGTGAAC AACTATAAGG	300
15	AAGAGCCCGT AATTGTAACC CCACCTACCA AACAATCACT AGTAAAAGTA CTAAAAACTC	360
	CAAAATGTGC ACTAAAATGG ATGATTTTGA GTGTGTACTC CTAAATTAGA ACACTTTGGT	420
	ATCTCTGAAT ATACTA	43€
	(2) INFORMATION FOR SEQ ID NO: 19:	
20	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 503 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:	
	TGTGCATATC CTGGGGAAAA AAATGGTACA TGTTTTAGAA ATTTTACTGT TTATAACAAT	60
	GCAGGCAGTC AGTTTCCCGT TTCAAACACA GATAGATACA TGCAACACTC AAGATCCTGC	120
	AGAGAGGCAG CCAGCATCTA TTGTTTAAAA AGGTTTCAAA AAGAATTCGG ATTGCTCKTT	180
	TCTCTTTTGA ATCTGTGTC CAAATGACAG GGACCAATAT TCGTCTTCTT TTTCKGTAAA	240
30	AYTCAGAAAG AMACATGAAA GAACCCAGAA TGCATTTCTT AAAGGGATTT AGTGCAGTTA	300

PCT/US98/04482

TTTTAAATAA TTTATGCACG CACACACAA TACATATATC CCCCGAGTAC ATATTTTTC	360
CCTTTTTACT TGTGTGCAAT CASTAGCTAC AATGACTGAA ATCCACTTCT TYGGGACTGI	420
GACATTTAAG CAAATCTTGT NTCTAGAAAN CGAAATGCCA NANTCTCGCA CAAAGCTGCT	480
CCGTCTGGGG CAACAAATCC ACA	503
(2) INFORMATION FOR SEQ ID NO: 20:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 358 base pairs	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: double	
(D) TOPOLOGY: linear	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:	
GGGCTGTCTC CCCAGTAGTA ACTTGCTGGC CCTGCCCTTG AAGTGGGGAA ACTGTGAAGG	60
GCTCCTTGAT CAAGCTTGTC CICTTTCTT ACCTCTTCCT CTCTTCTGTT TCCGCTGCAG	120
CTGAACAGGC CAGCAGGCAA CCTGCCATGG GGTCCTGCTC CAAGAACCGG TCCTTCTTCT	180
GGATGACTGG GCTCCTGGTA TYCATCAGCC TCCTCCTCAG TGAGTGGCAG GGTCCCTGGG	240
AAGGGAGGGC AATTGGAGAG GGCTGGGCTA GCTGGGCTCT GACCAACGGG TGGGCTGTTC	300
AACTTCTGAT GTCTTTGGGC AACAACACAG AAAAACACTC TGTTATGATT TACGAAAN	358
(2) INFORMATION FOR SEQ ID NO: 21:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 1926 base pairs	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: double	
(D) TOPOLOGY: linear	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:	
AGTGAAGGA GCTGGCCETG CEACTGGGCT TCGGGCCCTG TGCCAGAEGA GCANGCCTTC	6
CTGAGCAGGA GGAAGCAGGT GCTGGCCGCG GCCTTGAGGC AGGCCCTGCA GCTGGATGGA	120
GACCTGCAGG AGGATGAGAT CUCAGTGGTA GCTATTATGG CCACTGGTGG TGGGATCCGG	18
GCAATGACTT CCCTGTATGG GJAGGTGGCT GGCCTGAAGG AGCTGGGCCT CTTGGATTGC	24

	KTCTCCTACA	TCACOGGG 40	CICHGGGCTCC	ACTIGGGCCT	TYPODAACCT	TTATAAGGAC	300
	CCAGAGTGGI	A: @AA:0ADCD	CCRGGCAGGG	COCACTGAGI	DAEJAAGTETET	CCAGGTGACC	36(
	AAGAACAAGC	यक्षभग्रसम्बद्धाः वः	06/ AD 00004/A0	CAGCTGCAGC	ADEED! A TUD	GGAGCTGGCC	470
	GAGCGTGCCC	GCTTF FCTTA	COTOGRACOC	TOJACCIAACC	TOTOJOGOT	CATCAACGAG	48(
5	GOGOINGOIGO	ANJANGAGOC	CUATGATCAC	AAGCTCTCAG	AT CAACGGGA	GGCCCTGAGT	540
	CATGGCCAGA	ACCUTUTOOC	CATCTACTGT	GCCCCCAACA	CTAAAGGGCA	COAGCCCGACC	600
	ACTTTTGAAT	TTGGGGGAGTG	GTGCGAGTTC	TOTOCOTACG	AGGTGGGCTT	CCCCAAGTAC	6F.(
	G999CCTTCA	TYCCOCOTOTYSA	GETETTTGGC	TOCGAGTTOT	TTAT3G33CA	GCTGATGAAG	71(
	AGGCTTTCTG	AGTOCOGCAT	CTGCTTCTTA	GAAGGTATCT	G SAGDAADCT	GTATGCAGCC	78(
10	AACCTCCAGG	ACAGITTATA	СТЭЭЗССТСА	GAGCCCAGCC	AGTT TTGGGA	CCGCTGGGTC	840
	AGGAACCAGG	CCAACCTGGA	CAAGGAGCAG	GTCCCCCTTC	TSAASATAGA	AGAACCACC	900
	TCAACAGCCG	GCAGAATAGC	TGAGTTTTTC	ACCGATCTTC	TISACOTGGCG	TCCACTGGCC	96 (
	CAGGCCACAC	ATAATTTCCT	GCGTGGCCTC	CATTTCCACA	AAGACTACTT	TCAGCATCCT	1010
	CACTIVITCCA	CATYGGAAAGC	TACCACTCTG	GATGGGCTCC	CCAACCAGCT	GACACCCTCG	106:
15	GAGCCCCACC	TGTGCCTGCT	GJATGTNGGC	TACCTCATCA	ATACCAGCTG	CCTGCCCCT	114
	CTGCAGCCCA	CTOGGGACGT	GBACCTCATC	CTGTCATTGG	TODAACATCA	CCACGGAGCC	120
	TTCCAGCAGT	TGCAGCTCCT	GGGCCGGTTC	TGCCAGGAGG	DOOTAGEER A	GTTCCCACCC	1260
	ATCTCGCCCA	GCCCCGAAGA	GCAGCTCCAG	CCTCGGGAGT	GODACACCTT	CTCCGACCCC	1320
	ACCTGCCCG	GAGUUCCTGO	GRYGCTGCAC	TTTCCTCT\Э3	THEAGUGACTC	CTTCCGGGAG	1380
20	TACTCGGCCC	CTFFFFFFCCG	GCGGACACCC	GAGGAG3CG3	CAGCTGGGGA	GGTGAACCTG	144
	TCTTCATCGG	ACTOTOCOTA	CCACTACACG	AAGGTGACCT	ACAGCCAGGA	GGACGTGGA:	150
	AAGCTGCTGC	ACCTGACACA	TTACASTT	TGCAACAACC	AGRA ROAGCT	GCTGGAGGCT	156
	CTGCGCCAGG	CAGTGCAGCG	GAGGCGGCAG	CGCAGGCCCC	ACTGATGGCC	GGGGCCCCTG	162
	CCACCCCTAA	CTCTCATTCA	TTCCCTGGCT	GCTGAGTTGC	A PETPEGAAC	TGT CATCACG	168
25	CAGTGCTTNC	AGAGCCTCGG	GOTOAGGTGG	CACTGTCCCA	GESTOCAGGO	TGAGGGCTG3	174
	GAGCTCCCTT	GOGCCTCAGC	AGTTTGCAGT	GGGGTAAGGA	GGCCAAGCCC	ATTTGTGTAA	180
20	TCACCCAAAA	. cccccccgcc	TGTGCCTGTT	TTCCCTTCTG	CGITACCTTG	AGTAGTTGGA	186
30	GCACTTGATA	CATCACAGAC	TCATACAAAT	GTGAGGCGCT	GAGAAAAAA	AAAAAAAA	192
	ACTCGA						192

1224

E	(2) INFORMA	ATION FOR SE	EQ ID NO: 22	::			
5	(i)		HARACTERIST GTH: 1224 b E: nucleic	ase pairs			
10		(C) STR	E: NUCTEIC (ANDEDNESS: (OLOGY: line	dcuble			
	(xi) SEQUENCE I	DESCRIPTION	: SEQ ID NO	: 22:		
15	CCGCCGAAGC	TCCGTCCCGC	cceceeccee	CTCCGCCTCA	CCTCCCGGCC	GCGGCTGCCC	60
••	TCTGCCCGGG	TTGTCCAAGA	TGBAGGGCGC	TCCACCGGGG	TCGCTCGCCC	TCCGGCTCCT	12(
	GCTGTTCGTG	GCGCTACCCG	CCTCCGGCTG	GCTGACGACG	G3CGCCCCCG	AGCCGCCGCC	180
20	GCTGTCCGGA	GCCCCACAGG	ACGGCATCAG	ATTAATTAA	ACTACACTGA	AAGATGATGG	240
	GGACATATCT	AAACAGCAGG	TTGTTCTTAA	CATAACCTAT	GAGAGTGGAC	AGGTGTATGI	300
2.5	AAATGACTTA	CCTGTAAATA	GTGGTGTAAC	CCGAATAAGC	TGTCAGACTT	TGATAGTGAA	360
25	GAATGAAAAT	CTTGAAAATT	TGGAGGAAAA	AGAATATTT	GBAATTGTCA	GTGTAAGGAI	42:0
	TTTAGTTCAT	GAGTGGCCTA	TGACATCTGG	TTCCAGTTTG	CAACTAATTG	TCATTCAAGA	480
30	AGAGGTAGTA	GAGATTGATG	GAAAACAAGT	TCAGCAAAAG	GATGTCACTG	AAATTGATAT	540
	TTTAGTTAAG	AACCGGGGAAG	TACTCAGACA	TTCAAACTAT	ACCCTCCCTT	TGGAAGAAAG	€00
3.5	CATGCTCTAC	TCTATTTCTC	GAGACAGTGA	CATTTTATTT	ACCCTTCCTA	ACCTCTCCAA	660
55	AAAAGAAAGT	GTTAGTTCAC	TGCAAACCAC	TAGCCAGTAT	CTTATCAGGA	ATGTGGAAAC	720
	CACTGTAGAT	GAAGATGTTT	TACCTGGGCA	AGTTACCTGA	AACTCCTCTC	AGAGCAGAGC	780
40	CGCCATCTTC	ATATAAGETA	ATSTGTCAGT	GGATGGAAAA	GTTTAGAAAA	GATCTGTGTA	84(
	GGTTCTGGAG	CAACGTTTTC	CCAGTATTCT	TTCAGTTTTT	GAACATCATG	GTGGTTGGAA	900
45	TTACAGGAGC	AGCTGTGGTA	ATAACCATCT	TAAAGGTGTT	TTTCCCAGTT	TCTGAATACA	960
+ J	AAGGAATTCT	TCAGTTGGAT	AAAGTGGACG	TCATACCTGT	GACAGCTATC	AACTTATATC	1020
	CAGATGGTCC	AGAGAAAAGA	GCTGAAAACC	TTGAAGATAA	AACATGTATT	TAAAACGCCA	1080
50	TCTCATATCA	TGGACTCCGA	AGTAGCCTGT	TGCCTCCAAA	TTTGCCACTT	GAATATAATT	1140

TTCTTTAAAT CGTTAAGAAT CAGTTTATAC ACTAGAGAAA TTGCTAAACT CTAAGACTGC

CTGAAAATTG ACCTTTACAG TGCC

	(2) INFORMATION FOR SEQ ID NO: 23:	
5	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 694 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:	
10	GGCACGAGTC TTATTGTGCA CTGTAGCCTG AATCCCCCAG GGTAATTAAT ATGAAGTGCA	60
	AAAAGTTGAA TSTTUCAGTO TAAAAGGOAG TSSSAGAAAT TACATAGCAT GGAAATAATA	120
15	AAATBAACTO TIATTAATBA GAACGAGGCT CTTGCAGTGG CAAGTTCTGC TGGTCACCCC	180
	PETCHEDAAD OFFIBACACT DATAATIGGET TITTTITEDA ACTITCODES DETAGEGETA	240
20	GTACTTTTCA AAATSAGCTT GTTCTTCCTT CTSACACTCA TCTCAAAGCT CCATGGTGAC	300
20	GCAGAGGTOT GTTGAAGGTO ACAGGTCCTC GCTTGCATTG GCATACGGTC CTGTAGCATC	3€0
	ACTTETTAGC CLACTGCTGC TTGAGGAGC TAGGGGTATT CAGGGGATAGA GAGCTGAAAA	420
25	TAGGATTAAT TOOTTOOTTT TGACTOTOOC CTCAAGATGT COTTGCTTTG GTCTGAAAAAC	480
	ATCTCTTTAAC AACTTTTCAAAAAAAAAAAAAAAAAAA	540
20	TATTAGEATO TECCTOTO AGCOCTOGIA CERCECARA TETTETTET TETTETTOCA	6(1)
30	AGAGACTET TOTTECTOT TOACCOAGAA GTTTEAAACO AGOOTGEAA CATAGCAAGA	66.0
	CCCTATCTCT ACAAAAAAAA AAAAAAAAAA AAAA	694
35		
	(2) INFORMATION FOR SEQ ID NO: 24:	
40	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 796 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPCLOGY: linear	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:	
	ATGAGOGGO OTTPRRATISSO GOAGSTISSA GOSCOAA CAGGGGTCT GGGCCTGGCC	€.0
50	CTGCTGCTGCTGCCTCGGCCT CGGACTAGGC CTGGAGGCGC CGGGAGGCCCG CTTTCCACCT	120
	CGACCTCT3C CCAG3CCGCA CCCGAGCTCA GGCTCGTGCC CACCCACCAA GTTCCAGTGC	180
	COCACCAGIO GITTATIGOT GOCCOTOACO TIGIGOTIGO ACARGACTIO GACTIGOAGOS	240
55	ATGCCAGGGA TGAGGAGGAG TGCAGGATTG AGCCATGTAC CCAGAAAGGG CAATGCCCAG	300
	CGCCCCCTGG CCTCCCCTGC CCCTGCACCG GCGTCAGTGA CTGCTCTGGG GGAACTGACA	360
60	AGAAACTGCG CAACTGCAGC OGCCTGGCCT GCCTAGCAGS GRAGSKCMCG WKGCACGCTG	420

PCT/US98/04482

	AGGGATGACT GUATTCCACT CACGTGGCGC TGCGACGGCC ACCCAGACTG TCCCGACTCC	480
5	AGOGACGAGC TOWGOTGTEG AACCAATGAG ATOCTOTGG AAGGGGGATGC CACAACCATG	540
3	GGGCCCCTG TGACCCTGGA GAGTGTCACC TCTCTCAGGA ATGCCACAAC CATGGGGCCC	600
	CCMSTGACCO TOSSAGAGTET CCCCTCTGTC GSEBAATECCA CATCCTCCTC TGCCGEAGA	660
10	CAGTOTOGAA GOOCAACTEC CTATGOEGTT ATTECAECTG CTECGGTGCT CAGTECAGC	720
	CTGGTCACCG CLACCCTCCT CCTTTTGTCC TGGCTCCGAG CCCAGGAGCG CCTCCGCCCA	780
15	CTGGGGTTAC TGGTGG	796
20	(2) INFORMATION FOR SEQ ID NO: 25: (i) SEQUENCE CHARACTERISTICS:	
25	(A) LENGTH: 662 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ 1D NO: 25:	
30	TAATTOGGCA CGAGGCTCTG CTGGAGAAGG ACCTGCCCTG CCGCTGGGTT CTGAGCCGGA	60
2,0	GTGGTCG3TG G3T333ATGG AGGCGACCTT G5A3CAGCAC TTGGAAGACA CAAT3AAGAA	120
	TCCCTCCATT GTT33AGTCC TGTGCACAGA TTCACAAGGA CTTAATCTGG GTTGCCGCGC	180
35	GACCCTGTCA GATGAGCATG CTGGAGTGAT ATCTGTTCTA GCCCAGCAAG CAGCTAAGCI	240
	AACCTCTGAC CCCACTGATA TTCCTGTGGT GIGTCTAGAA TCAGATAATG GGAACATTAT	300
40	GATCCAGAAA CAGGATGGCA TCACGGTGGC AGTGCACAAA ATGGCCTCTT GATGCTCATA	360
	TOTGTTCTTC AGRAGOCTGT CATAGGAACT GRATICTACC TATGTTAATT ACCTTATAGA	420
	ACTACTAAAG TYTCAGTAGT TAGGCCATTC ATTTAATGTG CATTAGGCAC TTTTCTGTTT	480
45	ATTTAAGAGT CAATTECTTT CTAATGCTCT ATGGACCGAC TATCAAGATA TTAGTAAGAA	540
	AGGATCATGT TTT3AAGCAG CAGGTCCAGG TCACTTTGTA TATAGAATTT TGCTGTATTC	600
50	AATAAATCTG TTVGGAGGAA AAAAAAAAAAAAAAATTA CTGCGGNCCG ACAAGGGAAT	660
	TC	662
55	(2) INFORMATION FOR SEQ ID NO: 26:	

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1105 base pairs

(B) TYPE: nucleic acid

60

The Greenstand

	(C) STRANDEDNESS: double (D) TOFOLOGY: linear	
E	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26	
5	CCTGATGOTC TCTTTTCTGC AGTTCAAGGG AAAGACGAGA TCTTGCACAA GGCACTCTGC	60
	TYCTGCCCTT GGCTGGGGAA GGGT33CATG GAGCCTCTCC GGCTGCTCAT CTTACTCTT	120
10	GTCACAGAGC TGTCCGGAGC CCACAACAC ACAGTGTTCC AGGGCGTGCC GGGCCAGTCC	180
	CTGCAGGTGT CTTGCCCCTA TGACTCCATG AAGCACTGGG GGAGGCGCAA GGCCTGGTGC	240
	CGCCAGITGG GAGAGAAGGG CCCATGCCAG CITHITGGTCA GCACGCACAA CTTHITGGCTG	300
15	CTGTCCTICC TGAGGAGGTG GAATGGGAGC ACAGCCATCA CAGACGATAC CCTGGGTGGC	360
	ACTOTOACCA TTACGOTGCG GAATCIACAA COOCATGATG CGGGTCTCTA CLAGTGCCAG	420
20	AGCCTTCATG GLAGTGAGGC TGACACCCTC AGGAAGGTCC TGGTGGAGGT GTTCGCAGAC	480
	CCCCTGGATC ACCGGGATGC TGGAGATCTC TGGTTCCCCG GGGAGTCTGA GAGCTTCGAG	540
	GATGCCCATG TGGAGCACAG CATCTCCAGG AGCTCTTCKT AGGAAAGGCC GCAAATTCCC	600
25	ATTOCTICO CTCTTGCATA TCTTTCTCCT CCAAGACTG CATCTTTCTC ATCAAGATA	660
	TAGCAGCCAG CGCCCTCTGG GCTGCAGCCT GGCATGGACA GAAGCCAGGG ACACATCCAC	720
30	CCAGTGAACT GGACTGTGGC CATGACCCAG GGTATCAGCT CCAAACTCTG CCAGGGCTGA	780
	GAGACACOTG AAGGAAGATG ATGGGAGGAA AAGCCCAGGA GAAGTCCCAC CAGGGACCAG	840
2.5	CCCAGCCTGC ATACTTGCCA CTTGGCCACC AGGACTCCTT GTTCTGCTCT GGCAAGAGAC	900
35	TACTCTGCCT GAACACTGCT TCTCCTG3AC CCTGGAAGCA GGGACTGGTT GAGGGAGTGG	960
	AGAGOTOTA AGAACACTT ACAACTTTA DESTRUCTE ATATATA OESTODOO	1020
40	AATCCAAGAC TGTCATATTT AAAAAAAA AAAAAAAAA AAAFRRRRC CCCGGTACCC	1080
	AATTUGUUUT ATAGTGAGTU GTATA	1105
15		
45	AND THE STATE OF T	
	(2) INFORMATION FOR SEQ ID NO: 27	
50	(i) SEQUENCE CHARACTEFISTICS: (A) LENGTH: 1017 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:	
	· ····································	

CTCGCCTGGG CTGTTTCCCG GCTTCATTTC TCCCGACTCA GCTTCCCACC CTGGGCTTTC 60

CGAGGTGCTT TCGCCGCTGT CCCCACCACT GCAGCCATGA TCTCCTTAAC GGACACGCAG 120

WO 98/39446 PCT/US98/04482

177

	AAAATTGGAA	TGGGATTAAC	AGGATTTGGA	GTGTTTTTCC	TGTTCTTTGG	AATGATTCTC	180			
	TTTTTTGACA	AAGCACTACT	GGCTATTGGA	AATGTTTTAT	TTGTAGCCGG	CTTGGCTTTT	240			
5	GTAATTGGTT	TAGAAAGAAC	ATTCAGATTC	TTCTTCCAAA	AACATAAAAT	GAAAGCTACA	300			
	GGTTTTTTTC	TEGETEETET	ATTTGTAGTC	CTTATTGGTT	GGCCTTTGAT	AGGCATGATC	360			
10	TTCGAAATTT	ATGGATTTTT	Telettette	AGGGGCTTCT	TTCCTGTCGT	TGTTGGCTTT	420			
10	ATTAGAAGAG	TGCCAGTCCT	TYGGATYCCCTC	СТАААТТТАС	CTGGAATTAG	ATCATTTGTA	480			
	GATAAAGTTG	GAGAAAGCAA	CAATATGGTA	TAACAACAAG	TGAATTTGAA	GACTCATTTA	540			
15	AAATATTGTG	TTATTTATAA	AGTCATTTGA	AGAATATTCA	GCACAAAATT	AAATTACATG	60C			
	AAATAGCTTG	TAATGTICTT	TACAGGAGTT	TAAAACGTAT	AGCCTACAAA	GTACCAGCAG	660			
20	CAAATTAGCA	AAGAAGCAGT	GAAAACAGGC	TTCTACTCAA	GTGAACTAAG	AAGAAGTCAG	720			
20	CAAGCAAACT	GAGAGAGGTG	AAATCCATGT	TAATGATGCT	TAAGAAACTC	TTGAAGGCTA	780			
	TTTGTGTTGT	TTTTCCACAA	TGTGCGAAAC	TCAGCCATCC	TTAGAGAACT	GTGGTGCCTG	840			
25	TTTCTTTTCT	TTTTATTTTG	AAGGCTCAGG	AGCATCCATA	GGCATTTGCT	TTTTAGAAAT	900			
	GTCCACTGCA	ATGGCAAAAA	TATTTCCAGT	TGCACTGTAT	CTCTGGAAGT	GATGCATGAA	960			
30	TTCGATTGGA	TTGTGTCATT	TTATEAAATT	AAAACCAAGG	GAAACCCCAA	AAAAAA	1017			
,,										
35	(2) INFORMATION FOR SEQ ID NO: 28:									
) .)	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 391 base pairs									
		(B) TYP	E: nucleic	acid						
10			ANDEDNESS: O OLOGY: line							
	(xi) SEQUENCE I	DESCRIPTION	: SEQ ID NO	: 28:					
15	CCCTGGAAAG	AGGAACTGAT	GTTTGAG3GG	ACAGATGTGG	GTCACTTTCC	CTGGCAGTGC	€.C			
T.L.!	CCTCTAGCCT	TGCTGCCTTG	GCTTTCTGAC	CCCTTCCAGG	CTTCAGGGGC	CTGGGAGATC	120			
	TCATGCCTCA	GCCCAGGAAA	CATTTAATAG	GGAAAGCAGA	GACATGTCAT	GTCAGCCCCA	180			
50	CAGACAAGAA	TTTCTAGAGC	ACTIGTCCTG	TTGTTCCTTG	CCCCGACATT	ACTCAGTCTG	240			
	GGCCATGGAA	TCCATCCAAT	AAACACAGCA	ACACCCTATG	NTACTGACCA	AGCAAAGCTT	300			
55	GCCCCTGGTA	CCAAAGAGCT	AAATCATGAC	CAAAGTGTGA	CATGAATGTA	ACTGAAATGC	360			
<i>,</i>	GGGTTAGTTG	CTCAATGTAT	GCAAAGTCCC	A			391			

State of Live Consumer.

			7011			~ ~		C) C.
(シ)	TNECE	(M.A.)	11.5	FUR	5 m.C.	TD	NO:	29:

(i)	SEÇUENO	CE CHARAC	CTERISTICS:				
	(A)	LENGTH:	1139	base	pairs		

(E) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29.

10							
10	GGTGATATCT	TCATAGTGGG	CTATTACAGG	CAGGAAAATG	TTTTAACTGG	TTTACAAAAT	60
	CCATCAATAC	TTGTGTCATT	CCCTGTAAAA	GGCAGGAGAC	ATGTGATTAT	GATCAGGAAA.	120
15	CTGCACAAAA	TTATTGTTT	CAGCCCCCGT	GTTATTGTCC	TTTTGAACTG	TTTTTTTTT	180
	ATTAAAGCCA	AATTIGIGITTAA	GTATATATTC	GTATTCCATG	TGTTAGATGG	AAGCATTTCC	240
20	TATCCAGTGI	GAATAAAAAG	AACAGTTGTA	GTAAATTATT	ATAAAGCCGA	TGATATTTCA	300
20	TGGCAGGTTA	TTCTACCAAG	CTGTGCTTGT	TGGTTTTTCC	CATGACTGTA	TTGCTTTTAT	360
	AAATGTACAA	ATAGTTACTG	AAATGACGAG	ACCCTTGTTT	GCACAGCATT	AATAAGAACC	420
25	TTGATAGAA	CCATATTCTG	TTGACAGCCA	GCTCACAGTT	TOTTIGCCTGA	AGCTTGGTGC	48
	ACCCTCCAGT	GAGACACAAG	ATCTCTCTTT	TACCAAAGTT	GAGAACAGAG	CTGGTGGATT	54
20	TOATTAAT	CTTCGATATC	TGGCCATGGG	TAACCTCATT	GTAACTATCA	TCAGAATGGG	6(11
30	CAGAGATGAT	CTTGAAGTGT	CACATACACT	AAAGTCCAAA	CACTATGTCA	GATGGGGGTA.	66
	AAATCCATTA	AAGAACAGGA	TTAATAAAA	ATAAGATGAT	AAGCAAATGT	TTCAGCCCAA	72
35	TGTCAACCCA	GTTAAAAAAA	AAATTAATGC	TGTGTAAAAT	GGTTGAATTA	GTTTGCAAAC	78
	TATATAAAGA	CATATGCAGT	AAAAAGTCTG	TTAATGCACA	TCCTGTGGGA	ATGGAGTGTT	84
40	CTAACCAATT	GCCTTTTCTT	GTTATCTGAG	СТСТССТАТА	TTATCATACT	CAGATAACCA	90
40	AATTAAAAGA	ATTAGAATAT	GATTTTTAAT	ACACTTAACA	TTAAACTCTT	CTAACTTTCT	96
	TCTTTCTGTG	ATAATTCAGA	AGATAGTTAT	GGATCTTCAA	TGCCTCTGAG	TCATTGTTAT	102
45	AAAAAATCAG	TTATCACTAT	ACCATGCTAT	AGGAGACTGG	GCAAAACCTG	TACAATGACA	108
	ACCCTGGAAG	TIGCTTTITT	ATAAAAAAT	ATAAATTTCT	ТАААТСАААА	AAAAAAA	113

50

55

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 465 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30.

	CCACGCGTCC GCGGACGCGT GGGGAAGGTT TGTGCCAGTA GACATTATGT TACTAAATCA	6
5	GCACTITAAA ATCTTTGGTT CTCTAATTCA TATGAATTTG CTGTTTGCTC TAAITTCTTT	12
J	GGGCTCTTCT AATTTGAGTG GAGTACAATT TTGTTGTGAA ACAGTCCAGT GAAACTGTGC	180
	AGTAAAAAA AGTTAADADOOTT TTAADAAAAA AGTAAAAADOOA	24
10	TTACTGTCCA ACACAGTGGA GCAGCTTGTC CACAAATATA GTAATTACTA TTTATTGCTC	300
	TAAGGAAGAT TAAAAAAAGA TAGGGAAAAG GGGGAAACTT CTTTGAAAAAA TGAAACATCT	360
15	GTTACATTAA TGTCTAATTA TAAAATTTTA ATCCTTACTG CATTTCTTCT GTTCCTACAA	420
15	ATGTATTAAA CATTCAGTTT AACTGGTAAA AAAAAAAAA AAAAA	465
20	(2) INFORMATION FOR SEQ ID NO: 31:	
	(i) SEQUENCE CHARACTERISTICS:	
25	(A) LENGTH: 702 base pairs (B) TYPE: nucleic acid	
	(C) STRANDEDNESS: doubl∈ (D) TOPOLOGY: linear	
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:	
30	GCAACAAGCG GCCCACCTTC CTGAAGATCA AGAAGCCACT GTCGTACCGC AAGCCCATGG	60
	ACACGGACCT GGTGTACATC GAGAAGTCGC CCAACTACTG CGAGGAGGAC CCGGTGACCC	120
35	GCAGTGTGGG CACCCAGGGC CGCGCCTGCA ACAAGACGGC TCCCCAGGCC AGCGGCTGTG	180
	ACCTCATGTG CTGTGGGCGT GGCTACAACA CCCACCAGTA CGCCCGCGTG TGGCAGTGCA	240
40	ACTGTAAGTT CCACTGGTGC TGCTATGTCA AGTGCAACAC GTGCAGCGAG CGCACGGANG	300
10	ATGTACACGT GCAAGTGAGC CCCGTGTGCA CACCACCCTC CCGCTGCAAG TCAGATTGCT	360
	GGGAGGACTG GACCGTTTCC AAGCTGCGGG CTCCCTGGCA GGATGCTGAG CTTGTCTTTT	420
45	CTGCTGAGGA GGGTACTTTT CCTGGGTTTC CTGCAGGCAT CCGTGGGGGA AAAAAAATCT	480
	CTCAGAGNCC TCAACTATTC TGTTCCACAC CCAATGCTGS TYCACCCTCC CCCAGACACA	540
50	GCCCAGGTCC CTCCGCGGCT GGAGCGAAGC CTTCTGCAGC AFGAACTCTG GACCCCTGGC	600
	CCTCATCACA GCAATATTTA ACAATTTATT CCTGATAAAA ATAATATTAA TTTATTTAAT	66(
	TAAAAAGAAT TCTTCCAAAA AAAAAAAAAA AAAAAAACNI CG	702
55		

(2) INFORMATION FOR SEQ ID NO: 32:

60 (i) SEQUENCE CHARACTERISTICS:

A Section 18 Section 4.

(A) LENGTH: 1142 mase pairs

	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS doubl∈ (D) TOPOLOGY: linear	
5	(b) Topologi: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:	
	CGGCACGAGG AAGAAATGGC AGAGACIGGA ATCICTCTTC ATGAAAAAAT GCAGCCCTT	60
10	AACTTCAGTY CGACAGAGTG CAGCTCCTIC TCTCCACCCA CCACAGTGAI TCTCCTTATC	120
	CTGCTGTGCT TTGAGGGCCT GCTTTTCCTC ATTTTCACAT CAGTGATGTT TGGGACCCAG	180
15	GTGCACTCCA TCTGCACAGA TGAGACGGGA ATAGAACAAT TGAAAAAGGA AGAGAGAAGA	240
13	TG/GCTAAAA AAACAAAATG GATGAACATG AAAGCCGTTT TIGGCCACCC CTTCTCTCIA	300
	GG:TGGGGCA GCCCCTTTGC CACGCCAGAC CAAGGGAAGG CAGACCCGTA CCAGTATGTV	360
20	GTITGAAGGA CCCCGACCGG CATGGCCACT CAGACACAAG TCCACACCAC AGCACTACCG	420
	TOCCATCOST TOTCATGAAT GTTTAAATOG AAAAAGCAAA ACAACTACTO TYAAAACTTT	48(
25	TITIATETCT CAASTAAAAT GGCTGAGCAT TGCAGAGARA AAAAAAAGTC CCCACATTTI	540
23	ATTITIAAA AACCATCCTT TCGATTYCTT TYGGTGACCG AAGCTGCTCT CTTTTCCTTT	600
	TAAAATCACT TCTCTGGCCT CTGGTTTCTC TCTGCTGTCT GTCTGGCATG ACTAATGTAG	660
30	AGRECGETST CTCGCGCTGT GCCCATTCTA CTAACTGAGT GAGACATGAC GCTGTGCTGC	720
	GATGGAATAG TCTGGACACC TGCTGGGGGA TGCATGGGAA AGCCAGGAGG GCCCTGACCI	78
35	TCCCACTGOO CAGGAGGCAG TGGGGGGGCTC CCCGATGGGA CATAAAACCT CACCGAAGAT	84
33	GGATGCTTAC CCCTTGAGGC CTSAGAA94G CAGSATCAGA AGGGACCTTG GCACAGCGAC	90
	CTEATCCCCC AAGTGGACAC GSTITYSCCTG CTAACTCGCA AAGCAATTGC CTGCCTTGTA	96
40	CTTTATGGGC TYGGGGTGTG TACAATGATT TYGCGGGGGGA GTGGGGGGAGA AAGATGAAAG	102
	AGGTCTTATI TGLATTCTGA ATCAGGAATT ATATTCCCTG TGATTATTTT GAAGAGAGTGTG	108
45	TAGGALAGAC GTTTTTCCAG TYCAAAATGC CTTATACAAT CAAGAGGAAA AAAAAAAAAA	114
13	AG	114
50	(2) INFORMATION FOR SEQ ID NO: 33:	
	(i) SEQUENCE CHAFACTERISTICS:	
55	(A) LENGTH 928 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

WO 98/39446 PCT/US98/04482

181

	GGCACGAGGI CIAATGAGGG CICICITGII TGCTAGAGAI GAGAGAAAIG TATACIAAIC	60
	ATTIVATIT GTACTIAAAA TACATTIVAC TAATCATATI GATTIVAAT ATGACAAATI	120
5	CTTCTAGTAG ATACTAATCT ITCTTGTTTA TCATATTGTC CTAGAGAAGC CTAGGTAAAA	180
	ATGRETTCCA CCTAGTCTGT TTGTATACA CCTTCCCCCG TCCCCTCTCC ATCCCTGCCA	240
	ATTIGGGCTCT ATGCATATTG ACAAGCAAAT AAGAAAACCT TAGGTTCTTG TATTIGAATT	300
10	TCCAAAACAA TAAAAGGTTT TGACTCAAGA TTTGCATTCA AGAAGAGGCA GAAATTTTGT	36C
	CTIATCTYYT TATCATTYTG TGAACTTGTG TITCTCTGTA IGCTTAGAAA ATTTACACAC	420
15	AAGGAATGTT TGAAAAAGTG AGAATTTTAG AGTGCTTGGG TGGTTTTTAT TTGGTCAGTG	4 80
	CTGATGTGTT AGGTGTTTAG GGAAATAATG CTTCAGGACC TTTTTGACAA CACAGCTTCA	540
	TGAATGACTG GGGGATATTT ATGTTTGTGC TGAGAAAAGG GAGGGAGTGG GCAGGTTGGA	€00
20	GTGGGGACCT TTCCATTGAA AGCAGTGCAG TCAGCTGTTT CGTAGATGCA TTTTTTCTT	660
	ATGCTTGTAA CATTGTTCTT GTGTCCATAA TTGACTGAAA TGTCAAGCTC CAGGAATGCA	720
25	AGGCATTTAT CAGGTGACCA GAAGTAGAAC CTTGTTGATT ATGAAATGGA AGAATAATGT	780
	CAAGGTAGTG GGGGTAAAAT GACAAATAAG ATTTTACTGG TGAATTTCCA TGCTTAGTAT	840
	GTACATTAAC CTCTTTTTAA GTTGCATGTT AATCTGGTAT AACGTATTGT GTCTGGTTTA	900
30	TGCTTTGAGT AAAAAAAAA AAAAAAAA	928
35		
	(2) INFORMATION FOR SEQ ID NO: 34:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 773 base pairs	
40	(B) TYPE: nucleic aciâ (C) STRANDEDNESS: doubl€	
	(D) TOPOLOGY: linea:	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:	
43	GGCACGAGTT CIGGCCICIC ATTICCTIAC ACTCIGACAT GAATGAATTA TTATTATTI	60
	TOTTTYTOTT TYTTYTYTT ACATTTGTA TAGAAACAAA TTCATTTAAA CAAACTTATT	120
50	ATTATTATT TITACAAAAT ATATATATGG AGATGCTCCC TCCCCCTGTG AACCCCCAG	180
	TGCCCCCTG GGGCTGAGTC TGTGGGCCCA TTCGGCCAAG CTGGATTCTG TGTACCTAGI	240
55	ACACAGGCAT GACTGGGATC CCGTGTACCG AGTACACGAC CCAGGTATGT ACCAAGTAGG	300
55	The state of the s	360
	CARRETTIGGG CGCACCCACT GGGGCCAGGG GTCGGGGGAT GTTGGGAGCC TCCTCCCCAC	
	CARCCTTGGG CGCACCCACT GGGGCCAGGG GTCGGGGGGAT GITGGGAGCC TCCTGCCCAC CCCACCTCCC TCACTTCACT GCATTCCAGA TTGGACATGT TCCATAGCCT TGCTGGGGAA	420

عاميده المراجع المراجع المراجع

	AGGGGGCTGC TGGTGGGAAA TOYGAGCTAG GGCAGATGTA TGCATTCCTT TATGTCCCTG	540
5	TAAATGIGGG ACTACAAGAA GAGGAGCIGC CIGAGTGGTA CTVICTCTIC CTGGTAATCC	600
-	TOTOGOCOAS CONTATOSOA GARDAGANA TERANGANA ADROMATICO EADOCOGONO	€60
	TGGTCAAAAT CCCTGTGTAG CTGAATTCCC AAGCCCTGCA TTGTACAGCC CCCCACTCCC	720
10	CTCACCACCI AATAAAGGAA TAGTTAACAC TCAAAAAAAA AAAAAAAAA AAA.	773
15	(2) INFCRMATION FOR SEQ ID NO: 35:	
20	(i) SEQUENCE CHAFACTEFISTICS: (A) LENGTH: 453 base pairs (B) TYPE: nucleic acid (C) STRANDEDMESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:	
25	TAAAATGTTA CACGCTTGTC ATATTCCAGG CACTGCACTA TGTATGCCGT TTATCAACAG	60
	TTAGCTCAGC TAACCCTCAT GGTAACCTTG TTAGCCCCGA TTTTGCCAGA TGAGCAAAGT	120
20	GAGGTITITG AGGCCTIAAG TAACITGCCC AAGGTCACGT GGCTGKGAAG TAACTCTCCC	18(
30	AGTTCTGAGA 10000CGAGCC TH3ACGTTTT GTCATTGTAC ACCATCAACT CAGTGCTGCC	24(
	AGTCATTCCA GCAGCCAGCT AGCGTAGCICA AGGTTTCTCC ACCTTAGCAC TGTTGACATT	300
35	TCGAGCCAGA TAATTCTCTG TUUTGAUBAG CTGTCCTATG CCTTGTAGGA TATACAACAG	360
	CATCYTGGCT TTACCCACCA GATGYPHBAA CACCTCCCCA GTCGTGACAG CCCAAAATGI	420
40	CTATAGACGT TGCCACGTAT ACCCAGGGGT TCC	453
45	(2) INFORMATION FOR SEQ ID NO: 36:	
	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 459 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double 	
50	(D) TOPOLOGY: linea:	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3€:	
5	GTGACTGTCG CCCTGCCCGC AGCCALGTGG CCCCCGCTGT TGCTGCTGCT GCTGCTGCTC	60
55	COGGCOGCCC OGGTCCCCAC OGCCAAAGCC GCTCCCCACC CGGATGCTAA CACCCAGGAA	120
	GEOCTICAGA ACCIGCIOCA AFEASTOGGG GCTEGOGGAG ACGGAGAGACI GCGGECAGAC	180
60	TCACACCTEG CCCCGGGCTC TEGETGTATT GATEGGGCTG TEGTGECCAC GCGACCAGAA	240

WO 98/39446 PCT/US98/04482

	AGCCGGGGAG GAAGACCTGC GGTTCCGTGA GAGGCGTCCA GGGCTGCAGG CCACGGCGAC	300
4	AGGCTCCGGG GAACATGGGG CTTTCCCTGT CCACTCCCAA GGAGTGTGGG CCTCAACGCA	360
•	TTGGCAGGGG ACGGCCGTGT GCCCTCTYCA GACCCCCC CCAGATGCAT TYATTAGAAA	420
	TAAAAAAA TOOATTOTTT OTTAAATAAT	4 59
10		
	(6) 11707147771 707 070 77 17 17	
15	(2) INFORMATION FOR SEQ ID NO: 37:	
1.	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 509 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
20	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:	
	ATGAAATTTA CCACTCTCCT CTTCTTG3CA GCTGTAGCAG GGGCCCTGGT CTATGCTGAA	60
25	GATGCCTCCT CTGACTCGAC GGGTGCTGAT CCTGCCCAGG AAGCTGGGAC CTCTAAGCCT	120
	AATGAAGAGA TCTCAGGTCC AGCAGAACCA GCTTCACCCC CAGAGACAAC CACAACAGCC	180
30	CAGGAGACTT CGGCGGCAGC AGTTCAGGGG ACAGCCAAGG TCACCTCAAG CAGGCAGGAA	24(
	CTAAACCCCC TGAAATCCAT AGTG3AGAAA AGTATCTTAC TAACAGAACA AGCCCTTGCA	300
	AAAGCAGGAA AAGGAATGCA CG3AGGCGTG CCAGGTGGAA AACAATTCAT CGAAAATGGA	360
35	AGTGAATTTG CACAAAAATT ACTGAAGAAA TICAGTCTAT TAAAACCATG GGCATGAGAA	420
	GCTGAAAAGA ATGGGATCAT TGGACTTAAA GCCTTAAATA CCCTTGTAGC CCAGAGCTAT	480
40	TAAAACGAAA GCATCCAAAA AAAAAAAAA	509
40		
45	(2) INFORMATION FOR SEQ ID NO: 38:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 598 base pairs	
	(B) TYPE: nucleic acid (C) STRANLEDNESS: double	
50	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:	
55	ATGTTGGGCT GTGGGATCCC AGGGCTGGGC CTGCTCCTGC TGCTGCAGGG CTCGGCAGAC	6C
ب د	GGAAATGGAA TCCAGGGATT CTTCTACCCA TGGAGCTGTG AGGGTGACAT ATGGGACCGG	120
	GAGAGCTGTG GGGGCCAGGC GGCCATCGAT AGCCCCAACC TCTGCCTGCG TCTCCGGTGC	180
50	TECTIOCE ATERESTE PAINT PACIFICA CONTRACTO ENGRAPER	246

	ANGTOGOGO NEGICIGGAC GIGCAGOGGO CICCICCICO TGAGOTOLAG CATCIGCITG	300
5	TICTGETEES CCAAGCGCCE GGACETGCTG CATATGCCCE GTTTCCTEEC GGETCCGTGI	36(
2	GACATGTOCA AGTOCGTOTO GOTGOTOTOC AAGCACCGAG GGACCAAGAA GACGCCGTCC	42(
	ACGGGCAGOG TGCCAGTCGC CCTGTCCAAA GAGTCCAGGG ATGTGGAGGAG AGGCAGC	480
10	GGGGAAGGGA CGGAGGAGGG TGAGGAGACA GAGGGGCGAGG AAGAGGAGGA TTAGGGGGAGT	541
	CCCCGGGGGA CTGGTCAATA CAGATACGGT GGACGGAAAA AAAAAAAAA AAAAAAAA	59-
15		
	(2) INFORMATION FOR SEQ ID NO: 39:	
20	(i) SECATENCE CHARACTERISTICS: (A) LENGTH: 454 base pairs (B) TYPE: nucleic acid (C) STRANDEINESS: double (D) TOPOLOGY: linear	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:	
	ATGGAGGCTG TTTTTACAGT TTTTTTTTT GTTGTTGTTT TGTTTTTAAA GAATACAGAA	6(
30	GGAGCCAAGC TTTTTTGCAC TTTGTATCCA GCTGCAAGCT CAGGGCAGAG TCAAGGGCCT	120
30	OSSTTGGAAA AACCTGACTC ACAGGAATGC ATAATTGACC CTTGGAGCTA CCCAATAGCC	180
	CTTGGAGCTG GCACTGAACC AGGCTGCAAG ATTTGACTGC CTTAAAAACA CAAGGCCCTC	24(
35	TAGGCCTGGC AGGGATGTCC CTGTGCCCAG CACTGGGGGC TCGAAGACTG GTTTCTAGCA	30(
	CTACCGGTCA CGGCCATGTC GTCCTAGAAG GGTCCAGAAG ATTATTTTAC GTTGAGTCCA	36(
40	TTTTTAATGT TCTGATCACC TGACAGGCA CCCAAACCC CCAACTCCCA ATAAAAGCCG	42(
,,	TGACGTTCGG ACAAAAAAA AAAAAAAAA AAAA	454
45	(2) INFORMATION FOR SEQ ID NO: 40:	
50	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 425 base pairs(B) TYFE: nucleic acid(C) STRANDEDNESS: double(D) TOPOLOGY: linear	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:	
35	GCTAAAGGCC ATTCCCTCCG CAGGGCATTT GGCGTCGGGT GGGAGGGGAA AACGCATCTT	6(
	GTTAATTATT TYTAATCTTA TYTATTGTAC ATACCTGGGG CAGGGGGTTG GGGAGGTGGA	120
60	GGGGGRAGAA GGGTCCCCTC TCTCTGCCCC TCCCACTCCT TTTCTACGGC GATTTGTCTG	180

WO 98/39446

60

	TGTCTGGCCC CCACCCACTG MCCATCCCCC ATTGTTGTCT GGATGTGGTT CTATTTTTTA	240
5	TOGGTOTOOT THOUCOTOOT COUCOSTTYTO GOOCCOCGMOO CACCOCCTGC TOCCACTACC	30(
•	CTTTGTCTCT TGCTCTTTCT TGGGYTTCTG TACAACTCAA CTTGTATACA CTGTGTACAC	360
	ACAACCAGYC WAACGCAAAA COCAACGGCA AACACTTTAA AAAAAAAAA AAAAAACTGG	420
10	GGGGT	425
15	(2) INFORMATION FOR SEÇ ID NO: 41:	
20	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2471 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear 	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:	
25	GGCACGAGTA TGGCTTCCCG TGGACTCAGC CTCTTCCCCG ANTCCTGGCA CGAGGGGGCT	6(
	TCGCGTCTGT GCTTCCTGTG GCTGACGTCA TCTGGAGGAG ATTTGCTTTC TTTTTCTCCA	120
30	AAAGGGGAGG AAATTGAAAC ISAGTGGCCC ACGATGGGAA GAGGGGAAAAG CCCAGGGGTA	180
50	CAGGAGGCCT CTGGGTGAAG GCAGAGGCTA ACATGGGGTT CGGAGCGACC TTGGCCGTTG	24(
	GCCTGACCAT CTTTGTGCTG TCTGTCGTCA CTATCATCAT CTGCTTCACC TGCTCCTGCT	300
35	GCTGCCTTTA CAAGACGTGC CGCCGACCAC GTCCGGTTGT CACCACCACC ACATCCACCA	360
	CTGTGGTGCA TGCCCCTTAT CCTCAGCCTC CAAGTGTGCC GCCCAGCTAC CCTGGACCAA	42(
40	GCTACCAGGG CTACCACACC ATGCCGCCTC AGCCAGGGAT GCCAGGAGCA CCCTACCCAA	480
,,	TGCAGTACCC ACCACCTTAC CCAGCCCAGC CCATGGGCCC ACCGGGCCTAC CACGAGACCC	540
	TGGCTGGAGA GCAGCCGCCCCCCCCCCCCCCCCCCCCCC	600
45	GGATGCCCCG AAGGCGCCC TOTGAGCATT CCCTGGCCTC TOTGGGTGCC ACTTGGTTAT	660
	GTTGTGTGT TG00T3A9TG GT0T9CAGGC GCGGTTCCTT A0GCCCCATG TGTGCTGTGT	720
50	GTGTCCAGGC ACGETTCCTT ACGCCCCATG TGTGCTGTGT GTGTGCTGCC TGTATATGTG	780
20	GCTTCCTCTG ATGCTGACAA GGTGGGGACCAGAC AATCCTTGCC AGAGTGGGCT GGGACCAGAC	840
	TITGTTCTCT TCCTCACCTG AAATTATGCT TCCTAAAATC TCAAAGCAAA CTCAAAGAAT	960
55	GGGGTGGTGG GGGGCACCCT GTGAGGTGGC CCCTGAGAGG TGGGGGCCTC TCCAGGGCAC	960
	ATCTGGAGTT CTTCTCCAGC TTACCCTAGG GTGACCAAGT AGGGCCTGTC ACACCAGGGT	1020
	GGCGCAGCTT TCTGTGTGAT GUAGATGTGT CCTGGTTTCG GUAGCGTACC AGCTGCTGCT	1080

	TGAGGCCATG	ect cospece	CGGAGTTGGG	GGTACCCGTT	GUAGAGCCAG	GGACATGATY:	1140
	CAGGCGAAGT	TARRESTOR	GCCAAGTTGG	ACTITIGATICS	TTTYGGGCAGA	TGTCCCATTG	1200
5	CTCCCTGGAG	CCTGTGATGC	CIGTTGJGGA	TORGETARIO	TOOTGATGCO	AGAACACCTC	1260
	AGGCAGAGOC	CTACTCAGCT	GTACCTGTCT	GCCT9GACTG	TOCCOIGTOS	CCGCATCTCC	1320
10	CCTGGGACCA	ODDORAĐO ITO	CACATGGACA	CACAGCCTAG	CIGCCCCAG	GGAGCTCTGC	1380
10	TGCCCTTGCT	SCOOTGOCC	TTCCCACAGG	TGAGCAGGGC	TOCTGTCCAC	CAGCACACTY	1440
	AGTTCTCTTC	CCINCLAGIGT	TTTCATTTTA	TTTTAGCCAA	ACATTTTGCC	TGTTTTCTGT	1500
15	TTCAAACAT 3	ATAGTTSATA	TGAGACTGAA	ACCCCTG33T	TGTGGAGGGA	AATTGGCTCA	1560
	GAGATGGACA	ACCTGGJAAC	TGTGAGTCCC	TGCTTCCCGA	CACCAGCCTC	ATGGAATATG	162(
•	CAACAACTCI	TGTACCCCAG	TCCACGGTGT	TOTGGCAGGA	GGGACACCTG	GGCCAATGGG	1680
20	CCATCIGGAC	CAAA/9GT\ #93	GTGTGGGGCC	CT(+GATGGCA	GCTCTGGCCC	AGACATGAAT	1740
	ACCTCGTGTT	CCTCCTCCCT	CTATTACTGT	TTCACCAGAG	CTGTCTTAGC	TCAAATCTGT	1800
25	TGTGTTTCTG	AGTCTAGUGT	CTGTACACTT	GTTTATAATA	AATGCAATCG	TTTGGAAAAA	1860
	AAAAAAA	AAACTCGTAG	GGGGGGCCCG	TACCCAATGG	GCYCMMARAT	AGTAGARWAC	1920
20	FAAAAYAMCA	ANTIGOAACCA	AAGAGCGGCC	AG 30GANITT	TAAGAGGGCC	CCCTTTTGGU	1980
30	GGNATCCANT	TTAGCCGBGG	TTNITTAAGGG	: AAGTTGCNTG	G1G3GGGTTA	G3GCCCSGT	2040
	KYTWCTTCCA	ACCAAGGGTT	YTYGTGGTTA	. GGCCGGGTTG	G3CCCMATGG	GCTGGGCTG:	2100
35	GTAAAGTGGT	GGGTMA.YTGC	MATTGGGTAG	GGTGCTGCTS	GCATTCCTGG	CTGAGGCGG:	1160
	ATGGTGTGGT	AGCCCTGGTA	GCTTGGTCCA	GGGTAGCTGG	GOGGCACACT	TGGAGGCTGA	1220
40	GGATAAGGGG	CATUCACCCA	CAGTGGTGGA	. TGTGGTGGTG	GIGACAACCG	GACGTGGTCG	1280
40	GCGGCACGTC	TTGTAAAGGC	AGCAGCAGGA	GCAGGTGAAG	CAGATGATGA	TAGTGACGAC	1340
	AGACAGCACA	. AAGATGGTCC	AGCCAACGGC	CAA9GTCGCI	COGAACCCCA	TGTTAGCCTC	14(1(
45	TGCCTTCACC	CAGAGGCCTC	CTGTACCCCT	r GGGTTTTCCC	CTCTTCCCAT	CGTGGGCCAC	1460
	TCACTCGTGC	: C					2473

55

(2) INFORMATION FOR SEQ ID NO: 42:

(i) SEQUENCE CHAFACTERISTICS:

(A) LENGTH: 2659 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

	GGCACGAGCT	TTTCTCTAGA	GTCTGAAAGA	TECTAGAAAG	TTAAAATAAA	TAACTTACTT	61
5	AAGAGAATTA	TGGATCTTT	AAAATAATTA	ATIAACTTGA	TGATTTGAAC	TAACAGTTAT	120
5	GATAATTOTG	CTATTTATAG	CTTTTTTAT	TOCCOTGCAG	AAAACCATAG	GCAAAATTGC	180
	AACATGTTTG	GAATTGCGAA	GTGCAGCTTT	ACAGTCCACA	CAGTCTCAAG	AAGAATTTAA	240
10	ACTGGAGGAC	CTGAAGAAGC	TAGAACCAAT	COTAAAGAAT	ATTCTTACAT	ATAATAAAGA	300
	ATTCCCATTT	GATGTTCAGC	CTGTCCCATT	AAGAAGAATT	TTGGCACCTG	GTGAAGAAGA	360
15	GAATTTGGAA	TTTGAAGAAG	ATGAAGAAGA	GEFTGGTGCT	GGAGCAGGTC	TCCTGATTCT	420
15	TTCCTGCTAG	AGTTCCCGGT	ACTTTATTAC	CAAGGTTGCC	ATCGGAACCA	GGAATGACAT	480
	TACTCACTAT	CAGAATTGAG	AAAATTGGTT	TGAAAGATGC	TGGGCAGTGC	ATCGATCCCT	540
20	ATATTACAGT	TAGTGTAAAG	GATCTGAATG	GCATAGACTT	AACTCCTGTG	CAAGATACTY	600
	CIGIGGCITC	AAGAAAGAA	GATACATATG	TTCATTTTAA	TGTGGACATT	GAGCTCCAGA	660
25	AGCATGTTGA	AAAATTAACC	AAAGGTGCAG	CTATCTTCTT	TGAATTCAAA	CACTACAAGC	720
23	CTAAAAAAAG	GTTTACCAGC	ACCAAGTGTT	TIGCTTTCAT	GGAGATGGAT	GAAATTAAAC	780
	CTGGGCCAAT	TGTAATAGAA	CTATACAAGA	AACCCACTGA	CTTTAAAAGA	AAGAAATTG:	840
30	AATTATIGAC	CAAGAAACCA	CTTTATCTTC	ATCTACATCA	AACTTTGCAC	AAGGAATGAT	900
	CCTGACATGA	TGAACCTGGA	ACTTCTGTGA	ATTITACCAC	TCAGTAGAAA	CCATCATAGI	960
35	TCTGTGTAGC	ATATTCACCC	TTCAACAGGC	AGGAAGCAAG	CCGTACCCAG	ACCAGTAGG.	1010
55	CGGACGGAGT	CAAATGCAAA	GCTGTACCAC	AGAATTCAGA	GTCCAGCACA	TCACACTGAC	1080
	GTATAGGACT	CCTTGGGATA	CAGGITTATT	GTAGATTTTG	AAACATGTTT	TTACTTTTCT	1140
40	ATTAATTGTG	CAATTAATAG	TCTATTTTCT	AATTTACCAC	TACTCCTACC	СТ ССТ ТСС Т С	1200
	GAACAATACT	GTTGTGGGTA	GGATGTGCTC	ATCTTCAGAC	TTAATACAGC	AATAAGAAT G	1260
45	TGCTAGAGTT	TACACATCTG	TTCACTTTTG	CTCCAATATG	CTCTTTTGAC	TTAACGTCAA	1320
75	GCTTTGGGTT	GATGTGGGTA	GGGTAGTGTC	AAACTGCTTT	GAGAGGAATG	GGACCAGTT:	1380
	TGCTGCCTAA	GAAGGTCTGT	CTGGATGTTT	ATAGGCAGCA	CCTCTGAAGT	GGCCTAAATI	1440
50	CACCCTGATC	TGATAGTTTT	CCTGCTTAGA	AAGTGTGCCT	TGGCCAGATC	AGTATCCCAC	1500
	TOTOACECTA	TCCCTAGGTT	GTAGCTGTGA	TTGTTTCCAG	ATGACCAGAT	TGTTTTTCTG	1560
55	AAAATGAGCA	TATTTTAGT	CATGTCGATT	AGCTGTTCTT	CTACATCACA	TTGTTACTCI	1620
JJ	TTCTGATGAT	GATTCTAGGG	TTAACATTGG	AACCATCTCA	AAATAATTAC	AAAGTTTTAG	1680
	ATGGGTTAC	AATGTCTTCT	AAACAATGTA	ATCTAAAAT	AATTGAGTCA	GATGCTAAC:	1740
60	AGATACTGCA	GGCATAACTG	CIGITITICI	GACAACTGAT	TGTGAAACCT	TAAAACCTG	1800

PCT/US98/04482

188

	PTTOTODARA	TTACACTGAG	GASTATGCAA	AATCTYGAAA	GATATTOTAT	ATATTTTTTT	1860
5	TAGGTAGATA	GGATGGCCAT	ATOCTATALT	TTTAGATATA	CTGACATTCA	TCCATATGAA	1920
J.	AATATGCAGG	TCATTAGCTT	ACIATAATTT	ACTITITGACT	TAATGGGGCA	TAAATAAAAC	1980
	TTTCATAGTA	CACATGAGGT	GGATATITGA	TACACAGAAC	ATTTGCGTG	GGCTTTCTGT	2040
0	GGGTTAGATG	TAAAGCCCAC	TAATTTTATA	ATTCACTATT	TTAAATGAGO	AATGCATGAG	2100
	GGGAATGCAG	TGTCAGTACC	TGGCCTATTT	TTAAACTAGT	GTAATCACCC	TAGTCATACC	2160
15	ATTCAGTATG	TTTGCTTTTT	ALAATAAGTA	ACCACAATTA	AGTTGTTGTA	GCCCTTGCAC	1220
13	TTCAAGAGAT	CTAGTCTTTA	CTTTCAGTTG	TOTOTTAGGT	CCATTCTGTT	TACTAGACGG	1280
	ATGTTAATAA	AAACTATGCG	AGCCTGGAAT	GGAATTCTCC	AGCCAAATTT	TAGTCTTGTC	1340
20	CTCTCCATCT	TGATIGGATT	AATTCCAAAT	TCTAAAATGA	TTCAGTCCAC	AATAGCTCTA	2.400
	GGGGATGAAG	AATTTGCCTT	ACTTTGCCCA	GTTCCTAAGA	CTGTGAGTTG	TCAAATCCCT	1460
25	AGACTGTAAG	CTCTTCAAGG	AGCAAGAGGC	GCATTITCTC	CGTGTCATGT	AATTTTTCTA	2520
دی	AGGIGTTTGG	CAGCACTCTG	TACCCTGTGG	AGTACTCAGT	ACCITITIGIT	TGATGTTGCT	2580
	GACAAGACCT	GAAAAAAT	CCCTTAAAAA	AAAAACCCAT	TAAAGIGTAG	CAAAACCGAA	2640
30	AAAAAAWA	AAAAAAA					2659
35	(C) ThYC/CEM	מתוכאו ברום כ	FO ID NO. 4	٦ .			

(2) INFORMATION FOR SEQ ID NO: 43

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1635 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: doubl€

(D) TCPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

45	CGAGGAGGTC	ATGAACAAGG	AGGIGGGAGA	GETGGACGTG	GTGGCTATGA	CCATGGTGGC	60
	CGAGGGGGAG	GAAGAGAGAAA	TAAGCATCAA	GGAGGCTGGA	CAGATGGAGG	GAGTGGTGGA	120
50	GGAGGTGGCT	ACCANGATGG	TGGTTATCGA	GATTCAGGTT	TCCAGCCAGG	TGGCTATCAI	180
30	GETERCCACA	GCAGTGGTGG	CTATCAAGGC	GGAGGTTATG	GTGGCTTCCA	AACATOTTOT	240
	TCATATACAG	GAAGTGGATA	CCAG FRYGGY	GECTACCAGC	AGGACAATAG	ATACCAAGAT	300
55	GGCGGGCACC	ATYGGTGATCG	TYGGTYGGT	CETGETGGGC	GAGGTGGTCG	TGGAGGCCGA	360
	GGTGGTCGTG	CAGGCCAGGG	AGGAGGITGG	GGAGGAAGAG	GGA GCCAGAA	TTATCACCAA	420
40	GGGGTCAAT	TTGAACAGCA	TTTCCAGCAT	GJAGGTTATC	AGTATAATCA	TTCTGGATTT	480
60							

entropy of the secondary

	GGACAGGGAA GACATTACAC TAGTTGAGGO TAGCGAACCT TACATTTTGO TAGAGCTCAA	540
	GIAATAGAAA CTTAGTTTCA GAATOOTGAA TYCAGCACCT ATTTYGAATT AATGTGAGAC	600
5	CACAGETEGRA AGGERGAGATE CERTERIA ACCEPTERIO OFFICARES OCCUPANDA	660
	TYDESTATAT TYPETAGACI TACATAATGO CETTATTATTE TYPEAGAACA TACATATATATO	720
10	CTTTCTATGA AAAATTTTTT AAAAGSTSGT TAAAATTGCC TITAATSGC CASTAGACTA	780
10	ATTICACAGT CAGAACATGC AAACTTTTTT GAAGAAATTA CTTGAATAAG TAGTTTTCAT	840
	GTTTTCAATA TGCAGTTTTG AAAATSAGGA TTCACCTAGA CTTTTTTAGA TYTACTACYA	900
15	GGAAACCTTC CYCATATGAA TAACCATTTA TATGTGTTTT GCTTAAAGTA TTCCAATGCC	960
	TATTTTOCAA GCACAGTTCT GCCCCCCGT TGACTTTTAT GCCACGTGTG CTTCATGATG	1020
20	GAACTTTTAG GTCAGTTCCT ATTAAATGAG CTCTTYTCCA GATAGCACAT TCAGTAGCCT	1080
20	AADDE-BOADA ABTTOOAAAA BTOTOAADTO ETATACTATB TOATAABENA BTTBTTTTAT	1140
	AATCCATAAA GATTATAAAA GCAAACTAAG TTGTGAAGCT ATAGTACATG TAGGCATTA	1200
25	GTTAAGTATA GCAATTCAAA CTGACCTGCA TCCATCCAAA ACAAATTCCT CCTTCAACCT	1260
	TATTTTACT TGAAATTTGC TAGAAGAAAT AGCAAACCGA AATTTGTTTT ATGCATGAGT	1320
20	TAATACCACT GETTCAGCAA ATACAAGTTA GTTTGCTTTA AGCAGGTAAC TTTTTTTTA	1380
30	ATGGAAGAAA TGCACTACAA AGTTAAGACA GATTITTGCT AAGTGCAGGA GGCCCTTTAI	1440
	TATTGCTGCA GAAAACAAAA GCCTGGCTGA GTTGATGTTT TACATTCTCC CTTACTGAAA	1500
35	TCTACATGAC ATGATGCTTC TTGCTGGGTT TTTGTACATG TAAACATTGT CAAGCTGTGA	1560
	AAGAAAATGG CTGGAGGTGT GCTTTGTGTG AAAGGTGAGC ACTGAAAGTA TCTGTTAAGT	1620
10	TCTCCNGAAA AAAAA	1635
40		
45	(2) INFORMATION FOR SEQ ID NO: 44: (i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 780 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double	
50	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:	
55	AACATGGTCA TGTCTTTTAG TTTCATTATT TTCCTACTCC TTGTATGTCA AGAAATTACA	60
55	TTTTGCATGT CTTATGGAGA TGCTGTTAAT TGCTTCAGTG AGTGCTTTTC TAATCTGCAG	120
	ACCATTTACA TITCCTGTTT GCAGCATGCT GTGTGCAAAC AYTCAGTAAT TIGGAGTATT	180
60	CAATTATTTG TTAGGGCTCT TCCTATTTCC AAATGTGCTG AATTGTCTAT TGATGGGATT	240

	TOCABATOTT TTCATGAGAA CICGAAAATGI AGCTGGGTGG CACCIACCTA CGTTGCTAGG	30.
5	TACTCAQTAG ACTTTCTCTT COGLATAGIA AGCCLCAGAC AGITTLACT TITATCTACT	36
•	TIACITGIGG AAATAAAACA GICATTIIGI TOISAAAGAA TAAGAIAGGI TICTSIAGAG	427
	AAGSAATTCC TACCTCLAAA AGCTGCCTTG AGAACTCAGA ACTGGCAGTT TICTGAGGTG	48
10	ATTITITAAAT TICAGIATTA GOGAGAGICO AGCATTIGGI GACACAGATI CIACATAACI	5 4 0
	TEPRACACAL ACCEPTATE TO TOTOGTECHA ADMINAMENTAL ACCEPTAGE AT CTTGORAL ACCEPTAGE AT ACCEPT	60.
1.5	AGAACAAGTA GACTCTGGCA GCAGATCTCC AGAGACCCAA GTTTAGGTTC TCATAGTGTA	660
15	TYTGAAGTAG TINTACTCCI GGCTTAAGTA GTTTAGTGCC TGGGAGAATC CATTACTGAA	720
	AAGCATTTAA CTIAAAAAA AAAAAAAA AAAACTGAAA AGSTACTGAA TACAGAATAG	78(
20		
	ACCUMENTATION FOR CEC. ID NO. AL.	
	(2) INFORMATION FOR SEQ ID NO: 45:	
25	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2378 base pairs	
	(B) TYPE: nucleic acid (C) STRANDELWESS: double	
2.0	(D) TOPOLOGY: linear	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:	
	GOGAAGCAGO TGAAGCCGOO GOOGOGAGA ATTOCACGOTG GOTTOCCTGOG CCATGGTCAC	6(
35	CCACAGCAAG TTTCCCGCCG COGGGATGAG COGCCCCTG GACACCAGCC TGCGCCTCAA	12(
	GACCTTCASC TCCAAGAGOG ASTACCAGOT GEFEGTEAAC GLAGT-CECA ASTGCAGGAG	18(
40	AGCCGCTTCT ACTCGAGCGC AGTCACCCGC OSCCCAACCTGTTACT CAGTGCCGAG	240
40	CCCGCCGGCA CCTTTCTGAT COGCGACAGO TOGGGACCAG CGCCACTTCT TCACGCTCAG	300
	CGTCAAGACC CAGTCTG93A CCAAJAACCT GCGCATCCAG TGTJAAKG93G GCAGCTTCTC	360
45	TOTGCAGAGO GATOCCOGGA GOACHDAGOO OGTGSCOOGO TINGACTNGOG TGCTCAAGCT	42(
	GETGCACCAC TACATGCCGC COCCTUSAGC CCCCTCCTTC CCCTCGCCAC CTACTGAACC	481
50	CTCCTCCGAG GTGCCCGAGC AGCCGTCTGT CCAGCCACTC CCTGGJAGTC CCCCCAGAAG	540
50	AGCCTATTAC ATCTACTCOS GESCENCENAS AASACTATTAC STOTACTACTAC GESCONTON AASACTATTAC	6(-0
	CTOCAACGTG GOCACTOTTO AGCATOTOTA TOGGAAGAGO GOCAACGGCC ACOTGGACTO	660
55	CTATGAGAAA GTCACCCAGC TGCCGGGGGCC CATTCGGGAG TYCCTNGACC AGTACGATGC	71:0
	CCCGCTTTAA GGGGTAAAGG GCGTAAAGG CATGGGTCGG GAGAGGGGAC GCAGAGGCCCT	780
60	CTCCTCCGTG GCACATGGCA CAAGCACAAG AAGCCAACCA GGAGAGAGTC CTGTAGCTCT	84(
UU		

	GGGGGGAAAG	AGGGCGGACA	GGCCCCTCCC	TCTGCCCTCT	CCCTGCAGAA	TGTGGCAGGC	900
	GGACCTGGAA	TYSTGTTYGGAG	ADDEDDAAED	GTACCACCTG	AGTCTCCAGC	TTCTCCGGAG	960
5	GASCCAGCTG	TCCTGGTGGG	ACGATAGCAA	CCACAAGTG3	ATTOTOCTTC	AATTCCTCAG	1020
	CTTCCCCTCT	GCCTCCAAAC	AGGGGACACT	TCG3GAATG3	TGAACTAATG	AGAACTGCCA	1080
10	GGGAATCTTC	AAACTTTCCA	ACGGAACTTG	TTTGCTCTTT	GATTEGGTTT	AAACCTGAGC	1140
10	TGGTTGTGGA	GCCTGFGAAA	GGTGGAAGAG	AGAGAGGTCC	TGAGGGCCCC	AGGGCTGCGG	1200
	GCTGGCGAAG	GAAATGGTCA	CACCCCCCGC	CCACCCCAGG	CGAGGATCCT	GGTGACATGC	1260
15	TCCTCTCCCT	GGCTCCGGGG	AGAAGGGCTT	GGGGTGACCT	GAAAGGGAAC	CATCCTGGTG	1320
	CCCCACATCC	TCTCCTCCGG	GACAGTCACC	GAAAACACAG	GTTCCAAAGT	CTACCTGGTG	1380
20	CCTGAGAGCC	CAGGGCCCTT	CCTCCGTTTT	AAGGGGGAAG	CAACATTTGG	CACGAGATGG	1440
20	GCTGGTCAGC	TGGTCTCCTT	TTCCTACTCA	TACTATACCT	TCCTGTACCT	GGGTGGATGG	1500
	AGCGGGAGGA	TGGAGAGACG	GGACATCTTT	CACCTCAGGC	TCCTGGTAGA	GAATACAGGG	1560
25	GATTCTACTC	TGTGCCTCCT	GACTATGTCT	GGCTAAGAGA	TTCGCCTTAA	ATGCTCCCTG	1620
	TCCCATGGAG	A-3GGACCCAG	CATAGGAAAG	CCACATACTC	AGCCTGGATG	GGTGGAGAGG	1680
30	CTGAGGGACT	CACTGGAGGG	CACCAAGCCA	GCCCACAGCC	AGGGAAGTGG	GGAGGGGGC	174(
30	GGAAACCCAT	GCCTCCCAGC	TGAGCACTGG	GAATGTCAGC	CCAGTAAGTA	TTGGCCAGTC	1800
	AGGCGCCTCG	TGGTCAGAGC	AGAGCCACCA	GGTCCCACTG	CCCCGAGCCC	TGCACAGCCC	1860
35	TCCCTCCTGC	CTGGGTGGGG	GAGGCTGGAG	GTCATTGGAG	AGGCTGGACT	GCTGCCACCC	1920
	CGGGTGCTCC	CGCTCTGCCA	TAGCACTGAT	CAGTGACAAT	TTACAGGAAT	GTAGCAGCGA	1980
40	TGGAATTACC	TGGAACAGTT	TTTTGTTTTT	GTTTTTGTTT	TTGTTTTTGT	GGGGGGGC	2040
	AACTAAACAA	ACACAAAGTA	TTCTGTGTCA	GGTATTGGGC	TGGACAGGGC	AGTIGIGIGI	2100
	TGGGGTGGTT	TTTTTCTCTA	TTTTTTTTTT	TGTTTCTTGT	AATAATTTT	TGTTTACAAT	2160
45	CTGCCTCAAT	CACTCTGTCT	TTTATAAAGA	TTCCACTCCA	GTCCTCTCTC	CTCCCCCCTA	2220
	CTCAGGCCCT	TGAGGCTATT	AGGAGATGCT	TGAAGAACTC	AACAAAATCC	CAATCCAAGT	2280
50	CAAACTTTGC	ACATATTTAT	ATTTATATTC	AGAAAAGAAA	CATTICAGTA	ATTATAATA	2340
	AAGAGCACTA	TTTTTTAATG	AAAAAAAA	AAAAAAA			2378

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1772 base pairs

60 (B) TYPE: nucleic acid

(I) STRANDEDNESS: double (E) TOPCLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

5		•		-			
•	TOGATHOAHG	COTOCOUSG A G	GATOOVCAGO	AADDDTDDDD	GUCTGTGCCT	GAGCCTGAGC	€(
	CTGAGCTTUA	eenrekanne	GAGCOGATOG	CHAGGGGCTCC	GUUTTG 1966	ACCGCTGGGC	120
10	CCCCAGIGAT	G YTGACTCTYS	TOGGGAGGIC	TOTTICGGCT	TEETTCCITG	CTCAGCCTGI	180
	CGTGTCTGGC	GUTTTOCGTG	CTGCTGCTGG	CGCACTGTCA	GACGICGICA	AGAATTTCGA	240
15	GGATGTCAGA	ACECTAGA DEST	TOTGOCCTOC	CTATAAAGAA	AAATTCTGGG	CATATTTATA	300
13	ATAAGAACAT	ATTOTOAGAAA	GATTGONVATT	GCCTTCATGT	TOTOPSAGOOC	ATGCCTGTGC	360
	GGGGGCCTGA	TOTAGAAGCA	TACTGTCTAC	GETETGAATE	CAAATATGAA	GAAGAAGCI	420
20	CTGTGACAAT	CAASSTTACC	TTTAATATTA	ATCTOTOTAT	TTTVFGGCCTT	CTACTTCTGT	48(
	ACATGGTATA	TOTTMOTOTS	ACCYCADITED	TACTGAAGAG	GCGCCTCTTT	GGACATGCAC	54(
25	AGTTGATACA	GAGTGATGAT	GATATTUGUG	ATCACCAGCC	TYPTDGCAAAT	GCACACGATG	600
23	TGCTAGCCCG	GT CCCGCAGT	CGAGCCAACG	TGCTGAACAA	GGTAGAATAT	GGCACAGCAG	660
	CGCTGGAAGC	TTCAAGTCCA	AGAGCAGCGA	AAAGTCTGTC	TYTT\3ACCGGC	ATGTTGTCCT	721
30	CAGCTAATTG	AAETTAAEE®	TTCAAGGTGA	CTAGAAAGAA	ACARROAGAC	AACTGGAAA6	780
	GAACTBACTG	GGTTTTGCTG	GGTTTCATTT	TAATACCTTG	TIGATTTCAC	CAACTGTTG	840
35	TGGAAGATTC	AAAACTGGAA	GKAAAAACTT	GCTTGATTTT	THEFT	TAACGTAATA	900
55	ATAGAGACAT	TTYTAAAA90	ACACAGCTCA	AAGTCAGCCA	ATAAGTCTTT	TCCTATTTGT	961
	GACTTTTA-27	ATAAAAATA	AATCTGCCTG	TAAATAAAT	COTAAAAANT	TTTACCTGGA	102:
40	ACAAGCACTC	TCTTTTTCAC	CACATAGTTT	TAACTTGACT	TTYTEAAGATA	ATTTTCAGGG	1086
	THITIGITSI	TGTTSTTTT	TETTT ETTTG	TTTTGGTGGG	AGAPPGGAGG	GATGCCTGGG	114)
45	AAGTGGTTAA	CAACITTTTT	CAAGTCACTT	TACTAAACAA	ACTTTTGTAA	ATAGACCTTA	1907
73	CCTTCTATTT	TOGAGTTI JA	TTITATATITT	GCAGTGTAGC	CAGCCTCATC	AAAGAGCTGA	126
	CITACTCATI	TGACTITT GC	ACTGACIGTA	MATCTGGGT	ATCIGCTGTG	TCTGCACTI	1325
50	DOAAATEOTA	AAKATI TAGO	TIGCCTIGGTIGG	CTTTTCACAA	AAAGCAGATT	TTCTTCATGI	1384
	ACTYSTGATYST	CTGATGCAAT	GCATCCTAGA	ACAAACT33C	CATTTGCTAG	TTTACTCTAA	1440
55	AGACTAAACA	tagi cittggt	GTGTGTGGTC	TTACTCATCT	TCTAGTACCT	TTAAGGACAA	1500
JJ	ATCCTAAGGA	CTTGGACACT	TGCAATAAAG	TTATTTAAA	TTAAACCCAA	GCCTCCCTGG	1560
	ATAATAE/PTA	TATACACATT	TGTCAGCATT	TOOGGTOSTS	GTGAGAGGCA	GETGETTGAG	1620
60	CTCCAATGTG	TGCAGCTTTG	AACTAGGGCT	GGGGTTGTGG	GIGCCTCTTC	TGAAAGGTCT	1680

	PAGGAMATH GGATAACTGG CHTFFFFFT TCCTCTTTTAG PAGAAAAA TAAAAAAAA TAAAAAAAAA TAAAAAAAAA TAAAAAA	1740
5	TTTTGAAACA TCAAAAAAA AAAAAAAAA AA	1772
10	(2) INFORMATION FOR SEQ ID NO: 47. (i) SEQUENCE CHARACTERISTICS:	
15	(A) LENGTH: 1107 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: dcukle(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:	
20	CG99CGAGAA G99CAGACGG GACATGCAGC CTCTTCCGCC TGAGCCCCGG AAGTGATGTG	60
20	GCTGCGGCAT CGCGGCCTCG CTATGTCTGC CATTTTCAAT TTTCAGAGTC TATTGACTGI	120
	AATOTTGOTG CTTATATGTA CCTGTGCTTA TATTCGATCC TTGGCACCCA GCCTCCTGGA	180
25	CAGAAATAAA ACTYGATTGT TGGGTATATT TTGGAAGTGT GCCAGAATTG GTGAACGGAA	240
	GAGTOCTTO TOTOCAGTAT GCTGTATAGT AATGGCCTTC AGCATCCTCT TCATACAGTA	300
2.0	GCTGGGAAA ATGCCAGAAT GTAGTTGCCA TCAGATTTGA TTGTGAACAA GGACTGACTV-	360
30	CAGAAAATAA TGGAAAGGAT GTTTAACTCT TTTATCTCCG AACATTGAAT GAGATAAAT	470
	TCCAGATGCT GTTCTCTATT TTAATGTTAT TGGACCAGTG TTCTGTATAA ACAATTAAGA	48(
35	TGTAACCATT TAATAGTCTG TAACAATCAA CCTCAGTACT GTCACTACAA TATTACATTC	540
	TGCAAATETT ATTCTGTTGT ATCAGATACA AAATTTTAGT GAGGTATCTC TAAGGCACAT	600
	AGTAGAAAAC AAAATTGGTT AATTACTCAA GTTCCTTTCA CTGTGATTTG GAAATGATTT	660
40	AATCTTTATA GAATGAGAAC CTTTTTTGGA CTAGCTTTTT TATTAAAATG GCTCAATTTG	720
	TGTTGATAAG GATTGCATTA ATATTTAATA GTGCTTGCTT TTCCTCTGGG CACACCATTT	780
45	TGATCATTAA CCAGAGTACC TCTACTCTIA GCAAACTCTA GTTTATGACA AGTATTTAAA	840
	ATATTTAAAA CAAGCTTATG CAGTPOTTAA GBAGGAAGGT AAATBAGATG TAACTTAAAA	900
	ATAGTATIGG GAAAATGIIG ATAGIITAACA ITAGIGGAII TAGACIAGC AAAIGACAIA	960
50	GTAGGCTCTG AAACATCTTG TCAAGTATAT GIATTTTGTG CATGAATTTT TGCTGGAAAG	1010
	CTGTCTTTCT CTGAAAAACA CAACGTTCTT AGAATGAAAA GAACAATTAT AAAATAAAAA	
55	AAAAATTTAA AAAAAACTGG GCGGGGG	1107
))	WWWW TIWW WWWWCIAS ACASSA	/

Little Community

 $^{60\,}$ (2) information for SEQ ID NO: 48:

194

(i) SEQUENCE CHARACTERISTICS:

5	(A) LENGTH: 805 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TCPCLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEÇ ID NO: 4F:	
10	THECHAGAAGAG ATHEGATTHET THETTHEGAAAA CTACTACCGA TYGGCTGACG ATICTCTCAA	60
	THE TOTAL TOTAL ACTOR TO THE TENTER TO THE TENTER TO THE TOTAL TOT	120
	GBACAGCCAC CGAAACGTBA TGATGAGGTI GAATCTACAG CIGACCATGG GAACCTTCTK	180
15	TOTOTOGOTO TINGGACTAA DERSASTIGO TITTOGAATG AATTIGGAAT OTIOOOTIGA	240
	TACTICECTO ALDOCTACT STATTALEGA CATTACTICE TTTTTALEGA TACCAGGAGA	300
20	ISSTATIONTO DEPTACONDE AABANCHADA EGABENNOON TAOTHYDENO DEGERAREND	360
	ATGAAGGATA TGGTTCACGG CGGTATTCTG GAAGGGTTAT GATCATGGGC CCTAAAGTCA	420
25	CAGACTOTT TOCHTOAGOO TTOOCCCHAT CAGGGACACATTAG EGTOCOCGAG	480
25	TOUTTOAGIT TOUCTATOTG TOAGTTAAGT OGGTATTAOO TGCTTCATAG GGTTATGGGA	540
	AGAATTAAAC AATATETETA AAGCACTTA TAGCACACTO CCTAACACAA TAAGTTAGAA	600
30	ATATAATTIG TETAGAACIC TEACAACATA CATPIAAACA GATETIAETA ATTCIGGTAT	66(
	AAGSTITSTC ATAACCAAAT GBAAATGTAB GAAACATTTA TAATSTTETT AAAAGATAGA	720
35	AAATTOACCT CCATTTTCTT THETACTINGAA GATGGCACCA CINGGAATAAA TACTTAAGAC	780
22	ACTGAAAAA AAAAAAAA AACT	805
40	(2) INFORMATION FOR SEQ ID NO: 49:	
45	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1408 base pairs (B) TYPE: nucleic acid (C) STRANCEDNESS: double (D) TOPOLOGY: linear	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:	
50	TOATTATTTA TYCATGIGGC TGAAAGAGTA TATTAATTAT GITTAGATTT TYGGAAAAAG	60
	TOTGAACAAA AAAAGGACCT ATACAGTGCT CAAACTATAT TTTTAAAAAT ACTATTTTAT	120
55	TTTTACTCAC ATATGAAAAA AATGGCTGTA CIATCATGTT TACATACATA CTAACATTGG	180
	AAACAGAATA ACGAATTGTA TTTAAATTTT ATGAAGAACA CACAAACATT AAAACACTGA	240
60	TINGTIACAG AAAGCAGAST TIGAGGAAAA AACATTAGCT ATAATTITCA TITTIATTAA	300

WO 98/39446

195

AGAGCAGCAC CCTCTGAGAA TAATCAAACT GATTAGTAAT ATTCATCTAT ACTGCAAAAT 360

PCT/US98/04482

	MITTACOBAA COATTITOAT TATTITADIO ATOTTADICA TICAAAGGAA ACATOTATAA	42
5	ATGTTCCTCA CTCAATGCAA AGAAACTAAAA CATAATCTGA AGAAAAATAT GTCCTTATTA	48
	TIATTUACAA TAAAAAGTTG GCTTTATTCT GCAAGCCTGG GCATATTGTA CAATTGGCAG	54
10	TOTAABATAA DIDADOTADI OTTADITIDA ODAIBIAASI ADETDAASID BEDAAITOAD	60
10	OTTACOAAD TOAGATADO TOGAGEGTAD OTOAACOAGA OOAGTGOTO ATACOTACIO	66
	OTOAACDATA TAAAADAA DATOODADAA ADAATAAATO AOTOTADATO OOTATODIDT	72
15	ATCTATCTAT AAATTACATC TATATGCIAG CTCTTTAGTA TAAGTTGGAA AAAGGGGCCC	78
	TTTCTYGAGC ACATGGATAA AAGIATTATT GTAGTCTAAA GATTGCTGGA TTGATATTGT	84
20	GTTGTTATAA TGAAGATAAG CTACACACTG AAACCACTGT CAGATTAAGA AACTTCCACA	90
20	ACTTGTCTCA GTTCTTCAAA CAATGAAGCA AGTTCCTTTT CTAGGCTGAC AATTAGTCCT	96
	CTATTGGCAC TGCTGCTGGC TATGAAACTC ACCACCAAAG GTAAACGATT AAATTGAACC	1020
25	ACCTGSTAGG TGTTATAGTA ACAGATGATA CTTTTATTTT TGGAAAGTCC AAGTTTGCTT	1086
	CCTTGGTCTG TTGCAAGGC AAAAGTGJAT AAGAAACCAG GTCGCAAAAGC ATGCTCTGGA	1140
30	GCATTGTCAT TYGCCACTTT AATAACAGGT ACTCCATCTC TATCTGACAC AACAATGGCA	1200
30	TGGAGCCCTT CAACACTYGG TAACTYTYTA TACAAGAATC GCTTTAGGTC ATCCGCCATG	1260
	ATGAACCCCC TTCTCTCGCA GGATCAATCT CCACGCCTGG GGTTTCTGGG CTGCCTGGTT	1320
35	CTCTCCGCTG TCACTTCAGG GACAGCTTTA AAGACAGGTT CCTCCTCAAG CCACCGTCAC	1380
	ATGATTCATG ACCTCGTCTG CGCTCCA	1408
40		
10	(2) INFORMATION FOR SEQ ID NO: 50:	
	(i) SEQUENCE CHARACTERISTICS:	
45	(A) LENGTH: 1813 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
**	(D) TOPOLOGY: linear	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:	
	CATGGTGGGG CACGAGATGG CCTCTRACTC TTCWAACACT TCACTGCCAT TCTCAAACAT	60
55	GEGAAATCCA ATGAACACCA CACAGTIAGG GAAATCACTT TTTCAGTGGC AGETGAAGCA	120
	GGAAGAAAGC AAATTGGCAA ATATTTCCCA AGACCAGTTT CTYTTCAAAGG ATGCAGATGG	180
	TGACACGTTC CTTCATATTG CTGTTGCCCA AGGGAGAAGG GCACTTTCCT ATGTTCTTGC	240
60	AAGAAAGATG AATGCACTTC ACATGCTYGGA TATTAAAGAG CACAATGGAC AGAGTGCCTT	300

The State of the Comment of

	ICAGRIGGCA	ACCORTOPORD	ASCAGOASCS	CATTGTGTAG	GATICT/GGTGA	ACATOGGGGC	361
4	ACAGGTGAAC	ACCACAGACC	GCTGGYGAAG	AACACCTOTG	CATGIGIGIG	CTGAGAAGGG	42.
?	CCA TECTCAG	GTGIMTCAGG	CGATTCAGAA	GGGAGCAGTG	GGAAGTAATC	AGTTIGTGGA	480
	TOTT BARGOA	ACTAACTATS	ATGGCCTGAC	CACTTOCOOT	TYTYGCAGTCA	TAGCCCCACAA	541
10	TGCTGTGGTC	CAT GAACTCC	AGAGAAATCA	ACAGCCCAT	TOACCTGAAG	TOTAGGAGCT	600
	TITACTGAAG	AATAAGAGTI	TATASTTERT	CATTAAGTGC	CTAATT LAAA	TGGGAGCAGT	660
1.6	GGTVGGAAGCG	ADGOTAGGA	AAAGT3G003	CACAGCCCTG	CATTIGGCAG	CTGAAGAAGC	720
15	AAATCTGGAA	CTCATTCGOO	T CTTTTTYGGA	GOTGOCCAGT	TGCCTGTCTT	TTGTGAATGC	78:
	AAA/GCTTAC	ADIACOTAA	ADDTENESTO	NETTECTOCC	AGCTTGJAGT	ATCGGTTGAC	840
20	ACAATTAGAT	GCTGTCCGCC	TGTTGATGAG	GAAG3GAGCA	GACCCAAGTA	CTCGGAACTT	900
	GGA JAACGAA	CAGCCAGTGC	ATTINGGTICC	CGATGGCCCT	CTGGGA-JAAC	AGATCCGACG	96(
25	TATCCTGAAG	GGAAAGTCCA	TTCAGCAGAG	AGCTCCACCG	TATTAGETCC	ATTAGCTTGG	1020
25	AGCCIGGCTA	GCAACACTCA	. CTGTTAGTTA	GGCAGTCCTG	ATGTATCTGT	ACAT'AGACCA	1080
	TTTGCCTTAT	ATT GGCAAAT	GIAAGTTGTT	TOTATGAAAC	AAACATATIT	AGTICACTAT	114(
30	TATATAGTGG	GTTATATTAA	AAGAAAAGAA	. RAAAAATATC	TAATTWITCT	TGGCAGATTT	1200
	GCATATTTCA	TACCCAGGIA	. TCTGGATCTA	GACATCTGAA	. TTTGATCTCA	ATGGTAACAT	1260
35	TGCCTTCAAT	TAACAGTAGO	TTTTGAGTAG	GAJAGGACTT	' TGATTTETGO	CACAAAACAT	1310
ا ال	TATTAATAT	GCTATTGACA	GTTTCAAAGC	: AGGTAAATTG	TAAATGTTTC	TTTAAGAAAA	138(
	AG() ATGTGAA	AGGAAAAA(TAAATACAGO	TTGAGGCTI	CATTYGGCCI	TAGTCCCTGG	144
40	GAGTTACTGG	CGTTGGACAC	GCTTCAGTCA	A TTGGACTAGA	igaaaggtg1	CCATGGTTAG	150
	AATTTGATCT	TTGCAAACTC	OTTAATATAT	TIATTTTGI	T COUTAAAAAA	ATTIGTACATA	156
45	CTTGGTTGT	AACATGGTUA	A TATTYGAAAI	T GTATAAGTCC	: ATAAAATAGA	A AAAGAACAAG	162
-15	TGAATTGTT	G CTATTTAAA	A AAATYTTACA	A ATTOTTACT	A AGGAGTTTT	TATTGTGTAAT	168
	CA MAAGTO	I TIGIAGAIA	A AGCA/SATYSGY	GAGTTACGG/	GTTGTTCCT	TACTGGCTGA	174
50	AAGATATAT	I CGAATTGTA	A AGATISCTITI	r yctcatgcat	I TGAAATTAT	A CATTATTYGT	180
	AG-3GAATTG	C ATG					181

- (2) INFORMATION FOF SEQ ID NO: 51:
- (i) SEQUENCE CHARACTERISTICS:

 (A) LENGTH: 2070 base pairs

PCT/US98/04482

197

(B) TYPE: nucleic acid (C) STRANTEDNESS: double

(D) TYPCLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51: €(CCACGCGTCC GGAAGAGCGC GGCACTTCCG CTGGCCGGTG GCTGGCTGGC GGCTCCTGGA GGNGGTGGGG GGAGGGGAGG GGGGGGGGG CCCGGGGACT CGCATTCCCC GGTTCCCCCT 120 10 CCACCCCACG CGGTCTGGAC CATGGACGCC AGATGGTGGG CAGTGGTGGT GCTGGCTGCG 180 TTCCCCTCCC TAGGGGCAGG TGGGGAGACT CCCGAAGCCC CTCCGGATC ATGGACCCAG 240 15 CTATGGTTCT TCCGATTYST GGTGAATGCT GCTGGCTATG CCAGCTYTAT GGTACCAGGC 300 TACCTCCTGG TGCAGTACTT CAGGCGGAAG AACTACCTGG AGACCGGTAG GGGCCTCTGC 360 COCCOCTOS TGAAAGTTS TETETTTGC AATGAGCCA AGGCCTCTGA TGAGGTTCCC 420 20 CT9909CCCC GAACAGAGGC GGCAGAGACC ACCCGGATGT GGCAGGCCCT GAAGCTGCTC 48: TTOTGTGCCA CAGGGCTCCA GGTGTCTTAT CTGACTTGGG GTGTGCTGCA GGAAAGAGTG 5.40 25 ATBACCOCA GCTATGBBBC CACAGCCACA TCACCGGBTG AGCGCTTTAC GGACTCCACA 600 TTCCTGCTGC TAKINGAACCG AGTGCTGCA CTGATTGTGG CTGGCTCTC CTGTGTTCTC 660 720 TGEAAGCAGO CCCGGCATGE GGCACCCATG TACCGGTACT CCTTTTGCCA GCCTGTCCAA 30 784 GCTGGCCAAG GCCTCTAAGG TGATCCCTGT CATGCTGATG GGAAAGCTTG TGTCTCGGCG 840 35 900 CAGTAACGAA CACTGGGAGT ACCIGACAGC CACCCTCATC TOCATTGGGGG TCAGCATGTT TOTGOTATIC AGOGGACCAG AGRICOGGAG CTCCCCAGCC ACCACACTOT CAGGCCTCAT 46.0 1020 CTRACTGGCA GETTATATTG CPTTTGAACA GCTTCACCTC AAACTGGCAG GATGCCCTGT 40 TTGCCTATAA GATGTCATCG GTGCAGATGA TGTTTGGGGG TCAATTTCTT CTCCTGCCTC 1080 TTCACAGTGG GCTUACTGTT AGAAACAGGG GGCCCTACTG GAGGGAACCC GCTTCATGGG 1140 45 GOGACACAGI GAGITINGING COCANGECCI GETACTOTOC ATENGETOCS CATGIGGCCA 1200 GCTCTTCATC TITTACACCA TIGGGGCTT TGGGGCTGCC GTCTTCACCA TCATCATGAC 126 CCTCCGCCAS GCCTTTGTCA TACTTTTTTC CTGCCTTCTT TATGGCCACA CTGTCACTGT 132.0 50 GGTGGGAGGG CTGGGGGGTGG CTGHGGTCTT TGCTGCCCTC CTGCTCAGAG TCTACGCCGG 1380 1441 GGGCCGTCTA AAGCAACGGG GAAAGAAGCC TGTGCCTGTT GAGTCTCCTG TGCAGAAGGT 55 15.00 TTGAGGGTGG AAAGGGCTGT AGGGGTGAAATAGGA COCCCCACC ATCCCCTTCT GCTGTAACCT CTGAGGGAGT TGGCTGAAAG GGCAAAATGC AGGTGTTTTC TCAGTATCAC 1560 1620 AGACCAGOTO TGOAGCAGAG GATTYGGGAG CCCAGGAGGC AGCOTTCCCT TTTGCCTTAA 60

	GTCACCCATC TTCCACTAAS CASTTTATTC TGAGCCCSS GCTASACAG TCCTCAGTGA	1680
	GEGETTITIEG GEASTITINGS STORAGASAS CATA PETAGE TIPCACASTI ACTOTICOCA	1740
5	CAAGTFCCCI TAAGTCTTGC CCTAGCTSTG CTCTGCCACC TTCCAGACTC ACTCCCTCI	180-
	GCAAATACCT GCATTTCTTA CCCTGGTGAG AAAAGCACAA GCGGTGTAGG CTCCAATGCT	1861
10	OCCUTACION GASCALAGA ATSCISCISM TENTOCISMA GASCALAGA GASCALAGA GASCALAGA	1920
10	AGCACCACCA CCT:CTATE: TOCTGBATC CTAGECTCTS TTCCATGAGC CTGTTGCAG	1980
	TTTTGSTACI TTAGAAATGT AACTTTTCC TCTGATAATT TTATTTTAAA TTAAATTAAA	2040
15	ACTGCAAAAA AAAAAAAAA AAAAAAAAAA	2070
20	10) AUTONOMICON FOR OTHER TO MO. EG.	
20	(2) INFORMATION FOR SEÇ ID NO: 52:	
	(i) SEQUENCE CHAFACTEFISTICS: (A) LENGTH: 1416 base pairs	
25	(B) TYPE: nucleic acid	
25	(C) STRANDEDNESS: doubl∈ (D) TOPOLOGY: linear	
	(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 52:	
30	CCCTCACTAA AGGGAACAAA AGGTGGAGGT CCACGGGGT CGCGGCCGGCT CTAGAACTAG	€(
	TGGATCCCCC GGGCTGCAGG AATTCGGCAC ACGGATCCGC GTCCGCAGCG GGCCGCTGCT	120
35	GAGCTGCCTT GAGGTGCAGT GTTGGGGATC CAGAGCCATG TCGGACCTGC TACTACTGGG	18(
دد	CCTGATTGGG GGCCTGACTC TCTTACTGCT GCTGACGCTG CTGGCTTTTG CCGGGTACTY	24(
	AGGCTACTG GCTGGGCTGG AAGTGACTGC TGGGTCACCC CCCATCACCA ACGTCACTGT	300
40	GGCCTACAAG TICCACATGG GGCTCTATGG TGAGACTGGG CGGCTTTTCA CTGAGAGCTG	360
	CAGCATCTCT CCCAAGCTCC GCTCCATCGC TGTCTACTAT GACAACCCCC ACATGGTGCC	42 (
45	CCCTGATAAG TECCGATGTG CCETGGGCAG CATECTGAGT GAAGGTGAGG AATCGCCCTC	48(
4 .7	CCCTGAGCTC ATCGACCTCT ACCAGAAATT TGGCTTCAAG GTGTTCTCCT TCCCGGAACC	54(
	CAGCCATGIG GIGACAGOA COTTIOCCCI AACACCACCA TICIGICCCA ICIGGCIGGG	600
50	CTACCOGCOS TETECALCOT GEOTTGGACA COTACATOAA (GAAGEGAAAG CTGTGTGCCI	éti
	ATCCTCGGCT GGSGATCTAC CAGGAAGACC AGAATCCATT TCATGTGCCC ACTGGCACGG	720
55	CCAGGGAGAC TUCLATGIGO CTGAGATGAA GGAGACAGAG IGGAAAATGGC GGGGGGCTTGI	780
JJ	GGAGGCCATT GACACCCAGG TGGATGGCAC AGGAGTGAC ACAATGAGTG ACACGAGTTC	840

TGTAAGCTTG GAAGTGAGCC CTGGCAGCCG GGAGACTTCA GCTGCCACAC TGTCACCTGG

60 GGCGAGCAGC CGTGGCTGGG ATGACGGTGA CACCCGCAGC GAGCACAGCT AACAGCGAGT

900

96(

WO 98/39446 PCT/US98/04482

19¢

	CAGATTOCOGO DEDGAGO DESTITUACENT DELECATUTE TOTOCTOGOS CACCETEDACO	1020
5	GGAGTTACGS CTRGACCOTS GGACTTRAGC CCCTGGGGGA CTACCAAGTG GCTCTGGGAG	1080
-	CCCACTGCCC CTGAGAAGGG CAAGGAGTAA CCCATGGCCT GCACCCTCCT GCAGTGCAGT	114(
	TOCTGAGGAA CTBAGCAGAC TOTCCAGCAG ACTOTCCAGC COTOTTCCTC CTTCCTCTG:	1200
10	GGGAHGAGGG CTFCCTVEAGG GACCTGACTT CCCCTGCTTCC AGGCCTCTTG CTAAGCCTTV	1260
	TCCTCACTGC CCTTLAGGCT CCCAGGGCCCA GAGGAGCCAG GGACTATTTT CTGCACCAG:	1320
15	CCCCAGGGCT GCCGCCCTG TTGTGTCTTT TTTTCAGACT CACAGTGGAG CTTCCAGGAC	1380
1	CCAGAATAAA GCCAATGATT TACTTGTTAA AAAAAAAA AAAAAA	142€
20	(2) INFORMATION FOR SEQ ID NO: 53:	
	(i) SEQUENCE CHARACTERISTICS:	
25	(A) LENGTH: 1720 base pairs (E) TYPE: nucleic acid	
	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:	
30	GGCACGAGTG CGGCCCCAGC CTCTCCTCAC GCTCGCGCAG TCTCCGCCGC AGTCTCAGC1	ьÚ
	GCAGCTGCAG GACTGAGCCG TGCACCCGGA GGAGACCCCC GGAGGAGGGG ACAAACTTCG	120
35	CAGTGCCGCG ACCCAACCCC AGCCCTGGGT AGCCTGCAGC ATGGCCCAGC TGTTCCTGCC	180
	CCTGCTGGCA GCCCTGGTCC TGGCCCAGGC TCCTGCAGCT TTAGCAGATG TTCTGGAAGG	240
40	AGACAGCTCA GAGGACCGCG CTYTTCGCGT GCGCATCGCG GGCGACGCGC CACTGCAGGG	300
,,	CGTGCTCGGC GGCGCCCTCA CCATCCCTTG CCACGTCCAC TACCTGCGGC CACCGCCGAG	360
	CCGCCGGGCT GTGCTG39GT CTCCGCGGGT CAAGTGGACT TFCCTGTCCC GGGGCCGGGA	42(
1 5	GGCAGAAGTG CTG3T932GC GG3SAGTGCG CGTCAAGGTG AACGAGGCCT ACCGGTTCCG	48¢
	CGTGGCACTG CCTGCGTACC CAGGCGTCGT CACGGACGTC TGCCCTGGCG CTGAGCGAGC	540
50	TGCGCCCCAA CGACTCAGGT ATCTATCGCT GTGAGGTCCA GCACGGCATC GATGACAGCA	600
	GCGACGCTGT GSAGSTCAAG GTCAAAGSTA TCCCATCCAG ACCCCACGAG AGGCCTGTTA	660
	CGGAGACATG GATGGCTTCC CCGGGGTCCG GAACTATGGT GTGGTGGACC CGGATGACCI	7 <u>2</u> C
55	CTATGATGTG TACTGTTATG CTGAAGACCT AAATGGAGAA CTGTTCCTGG GTGACCCTCC	780
	AGAGAAGCTG ACATTIPGAGG AAGGACOGGC CTACTIGCCAG GAGGGGGGTG CAGAGATTTGC	84(

CACCACGGGC CAACTGTATG CAGCCTGGGA TGGTGGCCTG GACCACTGCA GCCCAGGGTG

60

The second of th

200

	PERRYTERTO TORDERSONA CODEASACTO TASSCOSATED SOTOTERSON TAUTSEATED	961.
	CTPSCCT99T GECAAGACTC TCTTCCTCTT CCCCAACAS ACTG9CTTCC CCAATAAGCA	102%
5	CAGCOSTTC AACSTOTACT GOTTCCGAGA CTORGCCCAG CTTOTGCCAT COOTGAGGC	1086
	TODAN TODAS COTTOAACOD ASCTITSATO GADIAGAGOD TATOSTCACA GISACASAGA	1140
10	CCCTGGAGGA ACTGCAGCTG CCTCAGGAAG CCACAGAGAG TGAATCCCGT GGGGCATCT	1200
10	ACTORATORO CATCATGRAG GACGERGAGE GAGALAGOTO CACTORAGAA GACCORGAA	1260
	AGGCCCTAG GANGCTCTA GAATTTGAAA CACAATCNAT GGTACCGCCC ACGGGGTTC	1320
15	CAGAABAGBA AGITAAGBCA TTGBAGBAAG AABAGAATA TGAAGATGAA GAAGAGAAA	1380
	AGGAGBABA AGAAGAGBAG GAGGTGBAGB ATGA VGCTCT GTGBBCATGG CCCAGCBAG	1440
20	TOAGCAGOOD GUICCOTGAG GCCTCTCTCCC CCACTGAGCC AGCAGCCCAG GAGGAGTCAC	1500
20	TOTOCIAGGI COCAGCAAGG GCAGTOCIGI AGUUTGITSC ATCACCACTT CCTGAIGGA'	1560
	ACTORDADA TYNOAGGOOT COAGGOOTA DOTGAGAOD TACTAGAGA CTOAGOOTACT	1620
25	CCAGGRAGAG GAACCTAGCA TCCCCATCAC CTTCCACTCT GGTTGAGGCA AGAGAGGTGC	1680
	GCGAGGCAC TYPTGGTCCT GAGCTATCTG GGTCCCTCGA	1720
30		
50	(2) INFCEMATION FOR SEQ ID NO: 54	
35	(i) SEQUENCE CHARACTERISTICS (A) LENGTH: 1117 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (B) TOPOLOGY: linea:	
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:	
	POPERFORM DEDERORACE ADDODED FOR DEATHERD BEOTHERD TANADAGE GROUND BEOTHERD TRANSPORMED ACTIONERS	€ (
	CGGGGAGICG AGGCATTTGC GCCTGGGCIT CGGAGGCGTAC CCAGGGGCCTG AGCCTTTGAA.	11:0
45	GCAGUAGGAG GGGAGGAGAG AGTGGGGCTC CTGTATCGGG ACCCCCTCCC CATGTGGAT	ĵ F (
	TOTOS GENEROS GENEROS AGRAPAS PO PRAGAZATO EN TOTOS GENEROS GE	240
50	GOTIGNOGGES CINENTIGOGG TOTOGETISTS OTHERSOGGES CONTIGUATE CATTY LAGIN	30(
	TOBAGATGEC TATGAACCT GTGTAAATGA AGGATGTGT GTTACCTACA ACAATGTG	36(
	AGGATACTIGO AAAGGTCCAG AAGGCTTCTT GFFFGAATAT TIGULAACATO GAGAULCCUK-	420
55		
	TGAGAAGAAC CGCTGCCAGA ATGGTGGGAC TYSTGTGGCC CAGGCCATGC TGGGAAAACC	480
	TGAGAAGAAC CGCTGCCAGA ATGGTGGGAC TIGTGTGGGC CAGGCCATGC TGGGGAAAGC CACGTGCCGA TGTGCCTCAG GGTTTACAGG AGAGGACTGC CAGTACTCGA CATCTCATCC	541

PCT/US98/04482

60

· . A. Wenner

	CTATGAGTGC ACCTGTCAAG TCGGGTTTAC ACSTAAGGAG TGCCAATGGA CCGATGCCTG	660
_	CONSTSTANCE COCTSTSCAA ATGGAAGIAS CHSTASCACT GTGGGGAACC ATTTGCTGCA	72C
5	AATGCCTCAC AGGCTTCACA GGGCAGAAGT GTGAGACTGA TGTGAATGAG TGTGACATT	780
	CAGGACACTG CCAGCATGGT GBCACCTGCC TCAACCTGCC TGGTTCCTAC CAGTGCCAG:	840
10	GCCTTCAGGG CTTCACAGGC CAGTACTGTG ACAGCCTGTA TGTGCCCTGT GCACCCTGC	900
	CTTGTGTCAA TGGAGGCACC TGTCGGCAGA CTGGTGACTT CACTTTTGAG TGCAACTGCC	960
1.5	TTCCAGAAAC AGTGAGAAGA GAAACAGCTTT TOAGAAAAC AGACAGGAA GTCTGGAAAC	1020
15	GAAAAGAACA CGATGAGAAT TAGACACTGG AAAATATGTA TGTGTGGTTA ATAAAGTGCT	1080
	TTAAACTGAA AAAAAAAA AAAAAAAAA AAAAAAAA	1117
20		
	(2) INFORMATION FOR SEQ ID NO: 55:	
25	(i) SEQUENCE CHARACTERISTICS:	
23	(A) LENGTH: 1903 base pairs (B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double (D) TOPOLOGY: linea:	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:	
	GGCACGAGCT CGGAGAGGCG GCGCCCCTGA GTAGGCCAGG AGCCTCTCTT GCAACTTCTG	60
35	CCACCGCGGG CCACCGCGGC CGCCTGATCC CGCAGAGGAA GGTCGCGGCC GTGGAGCGAT	120
33	GACCIGOGG GGTCCGGGC GGCGCCCGGG GCTGCCACAG CCGCCGCCGC TTCTGCTGCT	180
		240
40	GCTGCTGCTG CCGCTGTTGT TAGTCACCGC GGAGCCGCCG AAACCTGCAG GAGTCTACTA	300
	TGCAACTGCA TACTGGATGC CTGCTGAAAA GACAGTACAA GTCAAAAATG TAATGGACAA	360
	GAATGGGGAC GCCTATGGCT TTTACAATAA CTCTGTGAAA ACCACAGGCT GGGGCATCCT	
45	GGAGATCAGA GCTGGCTATG GCTCTCAAAC CCTGAGCAAT GAGATCATCA TGTTTGTGGC	420
	TGGCTTTYIG GAGGGTTACC TCATTGCCCC ACACATGAAT GACCACTACA CAAACCTCTA	480
50	CCCACAGGGTG ATCACGAAAC CTTCCATCAT GGATAAAGTG CAGGATTTTA TYGAGAAGCA	
	AGATAAGGTG GACCCGGAAA AATATCAAAG AATACAAGAC TGATTCATTT TGGAGACATA	
	CAGGOTATGT GATGGCACAA ATAGATGGOC TOTATGTAGG AGCAAAGAAG AGGGCTATAT	660
55		72C 780

GATGGGACAT GGGACATTGC TCCGCTCTTA TCAAGGTTCT TCCTGGATTT GAGAACATCC 840

	TYTTTTGTTCA	CTCAAGCTGG	TACACGTATG	CAGCCATGCT	CARGATATAT	AAACACTGGG	91,
	ACTICAACAT	CATAGATAAA	GATACCAGCA	STAGTOGOOT	CT CTTTCAGC	AGTTAGCCAG	96
5	GGTTTTTT93A	GTCICTGGAT	GATTITIACA	TTOTTAGOAG	ATABITTAERT	TTGCTGCAGA	102
	COACAAACAG	TAATTTETET	CECDOCAAAA	TAAAGCAGGT	AATADOOGAG	ACTITCCTGT	108
10	CCTGGCAAAG	ASTROCOTEA	GODAATATGA	TGGCAGATAG	TVGGCAAGAGG	ТЭЭЭС <mark>АGАС</mark> А	114
10	TOTTTTOAAA	TOTOAACATA	ATATOGAÇE	ACAATCAATA	CATGSTTCTG	GACCTGAAGA	120
	AAGTAAAGCT	GAACCACAGT	CTTGACAAAG	GCACTCTGTA	CATTGTGGAG	CAAATTCCTA	126
15	CATATGTAGA	AAETOTTATA	CAAACTGATG	AASQOATOTT	AGBATATTAG	CCCTCCTACA	132
	TTTCCTTETA	AAAAAETTADD	TOAAGACCC	GGAGTGGCTA	TOCACTOTTA	GTTCAGAAGC	138
20	ASSETTOSSEET	CTACICITAT	GATTTAGCTC	CACGAGCCAA	AATTTTCCGG	CGTGACCAAG	144
20	CAETEAAAED	TGATACGGCA	TAAAADTACCT	ATAT/CATGO	ATACAACAAT	TATAAGAAGG	150
	ATCCTTACAG	CAETTEDADAT	OCCUPATA TA	CCATCTGCTG	COSTGAGGAC	CCTGAACTCA	156
25	AACCCAATCC	GTCCTTGGAG	GTTETTATGA	CACAAAAGGT	GGCAGATATY	TACCTAGCAT	161
	CTCAGTACAC	COEFFATOOTA	OTEOTOAATA	CCACAGTACA	AGFTFFFCTC	COTGTTTTC	1€8
30	GCTGGGACCG	TTT-CAACAAA	ACTOTACATO	AGGCATGCC	AGAGGTCTAC	AA/CTTTGATT	174
,,,	TIATTACCAT	GAAACCAATT	TTGAAACTTG	ATATAAAATG	AA:93A:9G3AG	ATGACGGACT	180
	AGAAGACTGT	AAATAAGATA	CCAAAGGCAC	TATTTTAGCT	ATGITTTTCC	CATCAGAATT	186
35	ATGCAATAAA	TAATTATATA	TTGTCAAAAA	AAAAAAAA	AAA		190
40	(2) INFORMA	TION FOR SE	EQ ID NO: 56	5.			
			HARACTERIST:				
	(1)	(A) LEN	GTH: 1869 b	ase pairs			
15			E: nucleic ANDEDNESS:				
			CLCGY: line				
	(xi)	SEQUENCE I	DESCRIPTION	: SEQ ID NO	: 5e:		
50	ACAGCTTTTC	E A ECCCEPER	CACCACACCAC	CGAAGAGAGC	ASSENCESSO	CAAGCTCGAA	r)
	0700660060	CTC-CCCCTTC	CONFERENCE	CTCCCTCTGC	CCCCTCGGGG	TOBOGOGOOD	12
55	ACGATGCTGC	AGGGCCCEGG	CTC 3CTGCTG	CTGCTCTTCC	TOGGCTOGGA	CTGCTGCCTG	18
ن د	GGCTCGGCGC	GOGGGGTTT	COTOTTTGGO	CAGCCCGACT	TCTCCTACAA	GOGCANCAAT	24
	TIGICAAGCCCA	TOCCGGTCAA	COTGCAGOTG	TGCCACGGCA	TOBAATAGGA	GAACATGCG3	30

60 CTGCCCAACC TGCTGGGCCA CGAGACCATG AAGGAGGTGC TGGAGCAGGC CGGCGCTTGG

	ATCCCGCTGG	F TOATGAAGCA	GTGCCACCCG	GACACCAAGA	. AGTTCCTGTG	CTCGCTCTTC	42
5	GCCCCCGI CI	GCCTCGATGA	CCTAGACGAG	ACCATCCAGC	CATGCCACTC	GCTCTGCGTV-	48
J	CAGGTGAAGG	ACCGCTGCGC	CCCGGTCATG	TOOGCOTTOG	GYTTCCCCTG	GCCCGACATG	54
	CTTGAGTGTG	ACCETTTCCC	CCAGGACAAC	GACCTITIGCA	TCCCCCTCGC	TAGCAGCGA/	60
10	CACCINCINC	CAGCCACCGA	. GBAAGCTCCA	AAGSTATGTG	AAGCCTGCAA	CAAAAATAAA	660
	GATGATGACA	TAATACAECA .	GEAAACGCTT	TGTAAAAATG	ATTTTGCACT	GAAAATAAAA	720
15	GTGAAGGAGA	. TAACCTACAT	CAACCGAGAT	ACCAAAATCA	TCCTGGAGAC	CAAGAGCAAC	780
	ACCATTTACA	. AGITGAACGG	TGTGTCCGAA	AGGGACCTGA	AGAAATCGGT	GITGTGGCT	840
	AAAGACAGIT	TGCAGTGCAC	CTGTGAGGAG	ATGAACGACA	TCAACGCGCC	CTATCTGGTC	900
20	ATGGGACAGA	AACAGGGTGG	GGAGCTGGTG	ATCACCTCGG	TGAAGCGGTG	GCAGAAGGGC	960
	CAGAGAGAGT	TCAAGCGCAT	CTCCCGCAGC	ATCCGCAAGC	TGCAGTGCTA	GTCCCGGCAT	1020
25	CCTGATGGCT	CCGACAGGCC	TGCTCCAGAG	CACGGCTGAC	CATTTCTGCT	CCGGGATCTC	1080
	AGCTCCCGTT	CCCCAAGCAC	ACTCCTAGCT	GCTCCAGTCT	CAGCCTGGGC	AGCTTCCCCC	1140
	TGCCTTTTGC	ACGTTTGCAT	CCCCAGCATT	TCCTGAGTTA	TAAGGCCACA	GGAGTGGATA	1200
30	GCTGTTTYJA	CCTAAAGGAA	AAGCCCACCC	GAATCTTGTA	GAAATATTCA	AACTAATAAA	1260
	ATCATGAATA	TTTTTATGAA	GTTTAAAAAT	AGCTCACTTT	AAAGCTAGTT	TTGAATAGGT	1320
35	GCAACTGTGA	CTTGGGTCTG	GTTGGTTGTT	GTTIGTTGTT	TTGAGTCAGC	TGATTTTCAC	1380
	TTCCCACTGA	GETTGTCATA	ACATGCAAAT	TGCTTCAATT	TTCTCTGTGG	CCCAAACTTG	1440
	TGGGTCACAA	ACCCTGTTGA	GATAAAGCTG	GCTGTTATCT	CAACATCTTC	ATCAGCTCCA	1500
40	GACTGAGACT	CAGTGTCTAA	GTCTTACAAC	AATTCATCAT	TTTATACCTT	CAATGGGAAC	1560
	TTAAACTGTT	ACATGTATCA	CATTCCAGCI	ACAATACTTC	CATTTATTAG	AAGCACATTA	1620
45	ACCATTTCTA	TAGCATGATT	TCTTCAAGTA	AAAGGCAAAA	GATATAAATT	TTATAATTGA	1680
	CTTGAGTACT	TTAAGCCTTG	TTTAAAACAT	TTCTTACTTA	ACTITITGCAA	ATTAAACCCA	1740
	TIGTAGCTTA	CCTGTAATAT	ACATAGTAGT	TTACCTTTAA	AAGTTGTAAA	AATATTGCTT	1800
50	TAACCAACAC	TGTAAATATT	TCAGATAAAC	ATTATATTAT	TGTATATAAA	CTTTACATCC	1860
	TGTTTTACC						1869

Section 1989 August A.

- (2) INFOFMATION FOR SEQ ID NO: 57:
 - (i) SEQUENCE CHARACTERISTICS:
- 60 (A) LENGTH: 1259 base pair:

(B) TYPE: nucleic acid(C) STRANIEDNESS: double(D) TOPOLOGY: linea:

5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:	
	ACCEMENTOS TOGGEOGRADO GUSCUTOVAS CEVERAGORES ETGERIOTOS TETUROTOS	60
	GARCARAGO: OFFICACION GOSCOTTIFIO CACGATCCOG ARGRIGAGO CCGAGFFICO	120
10	COOPTIVIKUULA GYPOTTADI GARCATTIRRI GERAGRICCOOR GROGGOCORGI TRAVALTACTO	180
	ACCEPTARY CONTROL SOCIETATO CALECTRONO INCERCACION SOCIETANO	240
15	CGG: GGGGGGS NGGGCTCGGA GACAGCSTTT CTCCCGGAAAT CTTCCTCGGG CAGCARGTGG	300
	PODGECETAD DADGEOTERA DECONTROTT FORACECTOR DEGOGRACION MADELETERAD	360
20	GCAGCCCTGC GCGGGGCCCT GCTGGGCTGC CTCTGCCTGG CSTTGCTTTS CCTGGGCGC	420
20	COTOACCAOT TOSTTATUAST CAAAAAAACET PASTACOAAO AETECCETCOE CEAACAACACHOO	480
	COMPAGAÇÃO TATOCGAGAA AATMCAAAAC GACTETAGAG ACCOTOGGA TTACTOGACA	540
25	ATTTAACTTO COESTECTAD ATAATETAED AAESTEAAAAT AECCOESTAT CAESTACATA	600
	GAAGAGATTA AGGATCTITT GCCAGAAATG AGGGCATACT GGCCTGACGT AATTCACTOG	660
30	TTTCCCAATC GCAGCCGCTT CTGGAAGCAT GAGTGGGAAA AGCATGGGAC CTGCGCCGC	720
30	CARRINGATIS CRETICAACTO COARAARAR TARTITEGGOA GAAGOOTEGA ACTOTACAGE	780
	GAGETGEACO TOAACAGTGT GOTTOTAAAA TYGGGGATAA AACCATCCAT CAATTACTAC	840
35	CAAGTTECAG ATTTTAAAGA TGCCCTTECC AGAGTATATG GAGTGATACC CAAAATCCAG	900
	TGCCTTCCAC CAAGCCAGGA TGAGGAAGTA CAGACAATTG GTCAGATAGA ACTGTGCCTC	960
40	ACTAGGAGG ACCAGGAGCT GGAAAACTG: ACCAGGGG GGGAGCAGCC GTCCCCCAAG	1020
40	CABBAAGTOT GGCTGGCAAA TEBBGGCCGGC GAGAGCCGGG GTCTGAGAGT CTGTGAAGAT	1080
	GGCCCAGTCT TCTATCCCCC ACCIAAAAAG ACCAAGCATT GATGCCCAAG TTTTGGAAAT	1140
45	ATTCTGTTTT AAAAAGCAAG AGAAATICAC AAACTGCAGC TTTCTNAAAA AAAAANAAAA	1200
	AAAAATTGGG GGGTTTTTTT GGGGGGCCCCG GGGCCCTTGG TTTTTCCCCCC CGGGGGGG	1259
50		
50	(2) INFOFMATION FOR SEQ ID NO: 58:	
	(i) SEQUENCE CHAFACTERISTICS:	
55	(A) LENGTH: 1186 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double	

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

	CGGCATGGAG	AAMGGCTCCG	CTTCTGTTGC	AGCTGGCGGT	GCTCGGCGCG	GCGCTGGCGG	60
5	CCGCAGCCCT	CGTACTGATT	TOCATOGTTG	CATTTACAAC	TGCTACAAAA	ATGCCAGCA(121
2	TCCATCGACA	TGAAGAAGAG	TOTTOTTAAA	TAAATGCCAA	AGGCCAGAAA	GAAACTTTAC	18
	CCAGCATATG	GGACTICACCT	DAACAAADDA	TITCIGTOGT	TETECCTTCA	TACAATGAAG	241
10	AAAAACGGTT	GCCTGTGATG	ATGGATGAAG	CTCTGAGCTA	TOTAGAGAAG	AGACAGAAAC	304
	GAGATCCTGC	GTTCACTTAT	GAAGTGATAG	TAGTIGATGA	TGGCAGTAAA	GATCAGACCT	360
1.6	CAAAGGTAGC	TATAAATAT	TGCCAGAAAT	ATGGAAGTGA	CAAAGTACGT	GTGATAACCC	42(
15	TGGTGAAGAA	TCGTGGAAAA	GGTGGAGCGA	TTAGAATGGG	TATATTCAGT	TCTCGAGGAG	480
	AAAAGATCCT	TATGGCAGAT	GCTGATGGAG	CCACAAAGTT	TCCAGATGTT	GAGAAATTAG	54(
20	AAAAGGGGCT	AAATGATCTA	CAGCCTTGGC	CTAATCAAAT	GROTATAGIA	TGTGGATCT	€00
	GAGCTCATTT	AGAAAAAGAA	TCAATTGCTC	AGCGTTCTTA	CTTCCGTACT	CTTCTCATGT	€60
25	ATGGGTTCCA	CTTTCTGGTG	TGGTTCCTTT	GTGTCAAAGG	AATCAGGGAC	ACACAGTGTG	72C
23	GGTTCAAATT	ATTTACTCGA	GAAGCAGCTT	CACGGACGTT	TTCATCTCTA	CACGTTGAAC	78.
	GATGGGCATT	TGATGTAGAA	CTACTGTACA	TAGCACAGTT	CTTTAAAATT	CCAATAGCAG	٤ ٠,
30	AAATTGCTGT	CAACTGGACA	GAAATTGAAG	GTTCTAAATT	AGTTCCATTC	TGGAGCTGGI	900
	TACAAATGGG	TAAAGACCTA	STTTTTATAC	GACTTCGATA	TITGACTGGT	GCCTGGAGGC	960
35	TIGAGCAAAC	TCGGAAAATG	AATTAGGTTG	TTTGCAGTCT	TCAGTTGTGT	TCTTATGCT	102:
<i>J</i> .,	CAGTGTCACA	TTTCATTTCA	TTTGAAACTA	AAATTTTAAG	TAAAGCTGAA	ATAAACTTCT	1080
	TGTCATTGTC	TGCCTTTTGA	TAATTTTAAA	GAAATAACTT	TICATAAGTA	ATATTAAAAA	1140
40	TATCTCTTTG	GATATAAATG	ATTTTTAAAA	GATGTTTATT	TAAAAA		1184
45	(2) INFORM	ATION FOR SI	EQ ID NO: 5	9 :			
	(i)	SEÇUENCE C					
			GTH: 428 ba E: nucleic	-			
50		, -,	ANDEDNESS:				
		(D) TOP	CLOGY: line	ear			
	(xi) SEÇUENCE	LESCEIPTION	: SEQ ID NO	9: 59.		
55	GATCCCCCGG	CTGCAGGATT	CGGCACGAGT	ACTGATTCTT	CACTGAGCTT	KGTTAGTATA	60
	AGCAGAGTTC	CAAGTCTCCC	CTA 3CGTTGT	CTCTACATTT	CTTTATCATT	CCAGTGGGTA	12(
	RGGTTTAGCT	GCGGGAAGGA	CATTICATAA	GGGTTAGTTG	GACTGAGCAG	TATGGACATT	180

	TOCTTYTTTC ATTACTICATTATTC CTTCCTAGGAGGTTTGGT GGTTTAATA	24
	TIATTATTAT TOATGAGAG SUTGTGTGG SOCCEPAAA SECRETAGT TATTATTETY	3 (v.)
5	THISTOAR ISLAATERTO CURSALESS CRAISSAND STUDINTO TOTCATON NOTCATON	36:
	TYCAAAATAA AACTTTTTSA AATYTSAAAA AAAAAAAAA NAAAAAACTC GGGGGGGGG	411
10	COBSTACC	41)
10		
15	(2) INFORMATION FOR SEQ ID NO: 60:	
13	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 501 base pair:	
	(E) TYPE: rucleic ació (C) STRANDEINESS: doubl∈	
20	(D) TOPCLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60.	
25	GECACGAET TOTOLECATE ACAGECERA TRECEDENCA TOTOLECATE TOTOLECATE	€ (
20	TTGGTTTTCT GTGTGGGTCT CCTCACCATG GCCAAGGCAG AAAGTCCAAA GGAACACGA	120
	COSTTCACTT ACSACTACCA GEOCCEGGA ATCGSAGGC TCGTCATCGC CGGSATCCT	18.0
30	TTCATCCTSS GRATCCTCAT CGTSCTGAGC AGAAGATSCC GSTSCAASTT CAACCAGCAG	240
	CASAGGACTG GRRAACCCGA TGAAGAGGAG GRAACTTTCC GCAGCTCCAT CCGCCGTCT.	300
35	TOCACCOGRA GROGSTAGAA ACACCTRGARS OGATRGAATIC ORGOCAGRAC TOCCCTGGCA	361
55	OCTSACATOT COCACGOTOC AACTROCGCOCO COCACOGOCO COTOGOCACACO	410
	OCTOCOCO COTOCOCO ACCOMICAÇÃO ACCOMICAÇÃO DOCTOAGAÇÃO DOCOCOCOCOCO	481
40	A AAAAAATAAA AAAAAAAAA	50.
45	(2) INFORMATION FOR SEQ ID NO: 61:	
70		
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1197 base pair:	
50	(E) TYPE: nucleic acid (C) STRANTEINESS: double	
	(D) TOPCLCGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:	
55	ACATGATGEN TACCAAAGAA TTCGGCAAAAGA GCGCGCAGT GCAGCAGGTG CTCAATATCG	6.
	AGTGCCTGCG GGACTTCCTG ACGCCCCCGC TGCTGTCCGT GCGCTTCCGG TACGTGGGC:	124
60	CCCCCCAGGC CCTCACCCTG AAGCTCCCAG TGACCAKCAA CAAGTTCTTC CAGCCCACCG	187

PCT/US98/04482

غرفينيد ديا

	AGATGGCGGC CCAGGATTTC TICCAGCGCI GGRAGCAGCI GAGCCTCCCI CAACAGGAGG	240
	CECAGAAAAT CTTCAAAGCC AACCACCCA TGGACGCAGA AGTTACTAAG GCCAAGCTTC	300
5	ROCCERTOST DAADACTOOD AADOODACEN STAACAACETO OTOTOENONO EE TETECOEN	360
	CERESTOR COAGACTAAA GOOCTECAGE TOEGCTCT GOTTOGGCTE GAGCCCAAATS	420
10	COACTOOM OMECODEASO AACDACOO CETOCOADTO EDOCATORA ACCCEGACCO	480
10	TETETGAGET GETEGCACAG CASTTETGAG COCTEGACTO 16CCCCGGG GATGTFECC	540
	GCACTEBBOA GCCCCTTGBA CTBAGGCAET THIBETGBAT GBBBGACCTC CACTGGTGA	€00
15	AGABAAGACA CCAGGGTTTG GBBGATGCCT GBBACHTCC TCCGGCCTTT TGTATTTTTA	660
	TTTTTTTTA TCTGCTG TTTACATFOT G333GGTAG G33GAGTCCC CCTCCCTCCC	720
20	ADDOIDOOCT COTOTOTALED TRAASEEVACO SEERABAEER RABACURA DOGCOCONTT	78C
20	ADTRITATION TOTOCOCCO TOCOCOCC TOCCCATOCA GENERAT ATTATINGTE	840
	GOGAATAAAC AGAGAGACGC TAACAGCCCC ATGTCTGTGT CCATCACCCA CTGTTAGGTA	900
25	GTCAAAGAAG TGGGGTGAGG GCATGCAGAG TGTGGGTGGC CAGNTTCGCA GCCCATGGGT	960
	GGBACTCTGB GGAGACAGCA GCAGCAGCTAGCA ABGCCACCAG	1020
20	AGGIACTOI GTGCCTGGTT CCTYAGTCC CAACACCAGG TAGGAAGCTY TEGGIAGCTC	108€
30	GBOCTHBOTAG ACCTUATOTT CTGTOTTOTY THEFIGECOOT GCCTCTGGTG GGAAGTGCGT	1140
	GEAGGTEACC AGGETATAGA AGTTTCGEAS CTGATTGEAA GAEGATTAAC TTCCCC	1197
35		
40	(2) INFORMATION FOR SEQ ID NO: 62:	
40	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 595 base pairs	
	(B) TYPE: nucleic ació (C) STRANDEDNESS: double	
45	(D) TOPOLOGY: linear	
	(xi) SEÇUENCE DESCRIPTION: SEÇ ID NO: 62:	
50	ATTNANSACK TKYASCCTYT WATACMATCA TYATAGSSAR AAGCTSGTAC GCCTSMARGI	
50	ACCGGTTYGG AATTONCGGG TOGACCCAOG CGTCCGGCAC AGCGGGAGTT GGTTCTGACA	120
	CCAGATETTC TCTGCTCCTG GTTAATGTCA GTEAEGGETTG GAAGTTGAAT AAATEAGAAC	180
55	AGGAGTGGTC TGGGCCCATG TAAATGATCC TCCCTTGAAA GGAGGAACAG CTTTCATCAT	240
	TIGTTCCAGC TAAGCCTTGC ATGCATTATA GATCTGGTGC TAAGCAGTGG GAAAGATCTC	300
	ATAAGTAATG TITTATGTTC TITCTGTCTC TOSTCTTCTG TWSTTCTTGG CTTGTGGGTT	3€€
60	GIGTTTSTGT GTTAACTGGA AAATTGCTAT AAGCCAGTTG TCTCTAAGTT TTAAAAACGA	420

	ATTAGAAAAA CCATAAAATC TCTGGCCTAT GCACATTGTC CCTGTYYTGT GAAAACATTA	480
	NTTRABICIT ATALAAACTA CETETAATAA CTEACAACAB GAACEGAAAA TAAALEGGAAA	540
5	AAAAA AAAAAAAAA AAAACEAQA TOAAEAITAO DAACOADTAA TTAAAEAEAET	595
10	(2) INFORMATION FOR SEÇ IE NO: 63:	
15	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1478 base pairs (B) TYPE: nucleic acid (C) STRANDEDMESS: double (D: TOPCLOGY: linear	
•	(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 63:	
20	COGCICCIONAG GACGOACIOCOCICOCOCICOCATO AAGATOTOAA TITTGTGCGC	60
	AAGTTCCTAC AGCCCCTGTT GATTV-GAGAG CTGGCTCCGG AAGAACCCAG CCAGGATGGA	120
25	CCCCTGAATG CGCATGGTCG AGGACTTTCCG AGCCCTGCAC CAGGCAGGCGG AGGACATGAA	180
	GCTGTTTGAT GCCAGTGCCA CCTTGTTTGC TTTCCTACTG GGCCAGATCG TGGCCATGGA	240
30	GETACTGGCC IGGCTCCTTA TATACCTACT GGGTCCTGGC TGEGTGCCCA GTGCCCTGGN	300
50	CCG:CTTCAT CCTGGCCATC TUTUAGGTTC AGTCCTGGTG TCTGCAGCAT GACCTGGGCC	360
	ATGCTCCATC TTCAAGAAGW CCTGGTGGAA CCACGTGGCC CAGAAGTTCG TGATGGGGCA	420
35	GCTAAAGGGC TTCTCCGCCC ACTG (163AA CTTCCGCCAC TTCCAGCACC ACGCCAAGCC	480
	CAACATOTTO CACAAAGACO CAGACOTSAO GGSX53CGCC GTOTTOCTCC TY3GG53AGTC	540
40	ATCCGTCGAG TATGGCAAGA AGAAACGJAG ATACCTACCC TACAACCAGC AGCACCTGTA	600
40	CTTCTTCCTG ATCGCCCGC CGCTYCTCAC CCTYGTGAAC TTTGAAGTGG AAAATCTGGC	6 E C
	GUACATGODE GDETGOATGO A ELE RECEGA TOTEGOTOG GOOGGEGOT TOTATEGOOG	72(
45	CTTCTTCTTA TCCTACCTCC CCTTCTACGG CGTCCCTGGG GTGCTGCTT TCTTTGTTGC	780
	TGTCAGGGTC CTGGAAAGCC ACTGSTTCGT GIGGATCACA CAGATGAACC ACATCCCCAA	840
50	GGAGATOSGC CACGAGAAGC ACCGREACTG GGTCAGCTCT CAGCTGGCAG CCACCTGCAA	900
50	CGTGGAGCCC TCACTTTTCA CCAACTGSTT CAGCGGGCAC CTCAACTTCC AGATCGAGCA	96(
	CCACCTCTC COCCEGACACACACACACACACACACACCCCC COCCCCCCCCCC	1020
55	GCTGTGTGCC AAGCACGGCC TCAGCTACGA ATGAAGCCCT TCCTCACCGC GCTGGTGGAC	1080
	ATOGTCAGGT CCCTGAAGAA GTCTGGTGAC ATGTGGCTGG AGGCCTAGGT CCATCAGTGA	1140
	AGGCAACACC CAGGCGGGCA GAGAAGGGGT CAGGGCACCA GCAACCAAGC CAGCCCCGG	1200

PCT/US98/04482

	CCGGATCGAT ACCUUCACCO CTCCACTGGC CAGCCTGGGG GTGCCCTGCC TGCCTCCTG	1260
	OPENTODERVI ATTOCORACE TATATEMENTA DADITOCORE PONTOCONTO TEMPONTO	1320
5	TOTOGOGOTG ATGGGACAGO GOTAGAGGA AGGTGAGCAT AGCACATTTT COTAGAGCAA	1380
	GAATTSGSSG AAAGSTSTTA TITTTATATT AAAATACATT CAGATGTAAA AAAAAAAAA	1440
10	AAAAACTOGA GGGGGGCCC CGGNAACCAA TTCGCCCT	1478
15	(2) INFORMATION FOR SEÇ ID NO: 64:	
20	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 2033 base pairs(B) TYPE: nucleic acid(C) STRAMPEDNESS: double(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:	
25	GGCACGAGAA AGAACGCAAA GCTGAGAACA TEGACGTTAA TATCGCCCCA CTCCGCGCCT	60
23	GBGACGATIT CTTCCCGGGT TCCGATCGCT TTGCCCGGCC GGACTTCAGG GACATTTCCA	120
	AATGGAACAA CCGCGTAGTG AGCAACCTGC TCTATTACCA GACCAACTAC CTGGTGGTGG	18(
30	CIGCCATGAT GATTTOCATT GTGGGGTTTC TGAGTCCCTT CAACATGATC CTGGGAGGAA	240
	TOGTGSTGGT GOTGGTGTTC ACAGGGTTTTG TGTGGGCAGC CCACAATAAA GAOGTCCTTC	300
35	GCCGGATGAA GAAGCGCTAC CCCACGACGT TCGTTATGGT GGTCATGTTG GCGAGCTAT	360
Ju	TCCTTATCTC CATGTTTGGA GGAGTCATGG TCTTTGTGTT TGGCATTACT TTTCCTTTGC	420
	TSTTGATGTT TATCCATGCA TCGTTGAGAC TTCGGAACCT CAAGAACAAA CTGGAGAATA	480
40	AAATGGAAGG AATAGGTTTG AAGAGSACAC CGATGGCAT TGTCCTGGAT GCCCTAGAAC	54 0
	AGCAGGAAGA AGGCATCAAC AGACTCACTG ACTATATCAG CAAAGTGAAG GAATAAACAT	601
45	AACTTACCTG AGCTAGGGTT GEAGCAGAAA TYGAGTTGCA GETTGCCCTT GTCCAGACCT	6 6¢
7.	ATGTTCTGCT TGCGTTTTTG AAACAGGAGG TGCACGTACC ACCCAATTAT CTATGGCAGC	720
	ATGCATGTAT AGGCCGAACT ATTATCAGCT CTGATGTTTC AGAGAGAAGA CCTCAGAAAC	780
50	CGAAAGAAAA CCACCACCT CCIATTGTGT CTGAAGTTTC ACGTGTGTTT ATGAAATCTA	84 t.
	ATG3GAAATG GATCACACGA TTTCTTTAAG GGAATTAAAA AAAATAAAAG AATTACGCCT	900
55	TTTACAGCAA CAATACGATT ATCTTATAGG AAAAAAAAT CATTGTAAAG TATCAAGACA	960
5.1	ATACGAGTAA ATGAAAAGGC TGTTAAAGTA GATGACATCA TGTGTTAGCC TGTTCCTAAT	1026
	CCCCTAGAAT TGTAATGTGT GEGATATAAA TTAGTTTTTA TTATTCTCTT AAAAATCAAA	1080
60	GATGATCTCT ATCACTTIGC CACCUSTITG ATGIGGAGTG GAAACTGGTT AAGCCAGTTG	1140

	TINATACIO CITETINIA GEGATIVETO DA CALADACIA CALINITATIA	1266
5	ARTITITIAA AAGGUTGATU ATGAAGGA AGTITIGUGU AAUGAAGGA AGGITITITAA	12 6 %
-	AGAITTCCTTA AGAIYSTIAC AGTETATCO AATETATCT COTGGACGGA CTTATTAAAA	131'
	TESTITICITA DACETICDAS STATEATHAD ITOAAAANAA AASAAAAD ADADAADAS	1360
10	GTTACAGTGA AAAAAAT991 CCAAGAAAAT GTTTGCCATT TTTGCATTGT TTCGTTTTTA	144
	ACTORAGIA TIAGAAASAA GAAAAAGAAT GTOCATTIA TIAATICCTI AGRGCACAA	150:
15	AGAGGAGA TAGATTOT DOMAMAGET TAGAGGAGA TAGTOGATA AGAGGAGGAGA	1567
1.	AAGITTIAAA AAAIXSIAAI GAMAATSSAA TSCAGITACI SCASIIAATA AAAAATTTTE	1611
	GATAGCAATT GTTACAACCA TATGCCTTTA TAGCTAGACA TLAGAATTAT GATAGCATGA	3.68+
20	AAAAATAAT9 DATAAAATAIT TITEMAGTGT TTOOCTOVYT LIMADIAIGI TAGAIATHE	1740
	ATETOTTED ATTACAAAC ATGACTORA ACTOTACTA ACATTACOCE DOGSTTTENTA	1800
25	GUATTAAATT GYGAAGAUAA CTWGAGTGGT ACTTACTGAA GAAACTCTCT GTATGTCCTA	1860
	GAATAAGAAG CAACGATSTS CINCTICTGA TTTTTCTT3C ATTTTAAATT CFCAGCCAAC	1920
	CTACAGCCAT GATCTTTAG: ACAST SATAT CACCATGACT TCACAGACAT GGTCTAGAAT	198
30	AAA AAAAAATTEE AAATTAERTTA AAATAABAAE PAIADADDOA TTOODATERTO	2000
35	(2) INFORMATION FOR SEC ID NO: 65:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 440 base pairs (E) Type: nucleic acid	
40	(C) STRANCEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEC ID NO: 65	
45	ATGTTTCTTA CTAGAATAGT GEGTCCAAGC TATATAGCCC TAACTIFCCT GGTTTACATI	45.7
40		120
	GTGGCCCTAG TATCIGGGCA GCIGTGCATG GAGATAGCCA GAGGAAACAT TTTTTTTCTI	
50	AATGAATTGG TGACCACATT TTSTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	280
	CCT-GETTTCT TCTACAGIAS UTUATGTAAA TGTTGTTTTG UCCTTGTCGT TCTCAGTAGA	
5.5	ATTIGGTTCTG TAAACGAAAC CTGCTCCTGT AATTICAGTA TATGCTCATA TCTCATCTTT	300
55	GGCTCTCCCA TTTHCACAGAC AGUGATCCCT AAAAGATGTG CCCTAGAGGA TATCCAGAAC	
	AATCLAATTG GATGTCTTCI COGCTGCACT CCAGCTGCAGA AGACAGAGA AGACTCNATC	420
60	ТСАААААА ТТААААААА	44

الإختياء المراجع المراجع

•	(2) INFORMATION FOR SEQ ID NO: $\epsilon\epsilon$:	
5	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 3301 base pair: (B) TYPE: nucleic acid	
10	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:	
	GGTCATAAGG GGAGGGTTGN NGTGTGTCCC TCCAGGTTGT GCAGAGGGGA TTAGAAGTAA	6(
15	GTAGGTTAGA GGGGAGGTGG AGGGAGTGTG CTGGGGTGTG AGGTTTTATG ATGCTGAAAG	120
		180
	GATCATGATA TGCTAAGGAC AGGATAGTGT TGGGTTGTAC ACACAGGTGT AGGCAATCCT	
20	GGTGGCTAGT ATGLAAAGT GALTGTCTTA GTCCTTAG AGGGTACCTG NCAGAGTGCC	24(
	CTTGGARGGA CLAGTGCTGG AGAAATTAAT AGGAGAGGGG ACGGGCATCC ATTAACCTT.	300
	TCTTGCCTGC AGCCTGTAGG GTCCAGCGTC AAAGCGAATC ATGGGGTCCA GGGCTGAGCI	360
25	GTGCACTCTC TIAGGCGGAT TCTCCTTCCT CCTGCIACTG ATACCAGGCG AGGGGGCCAA	420
	GGGTGGATCC CTCAGAGAGA GTCAGGGAGT CTGCTCCAAG CAGACACTGG TGGTCCCGCT	48(
30	CCACTACAAC GAGTCCTACA GCCAACCAGT GTACAAGCCC TACCTGACCT TGTGCGCTGC	540
	GAGCGCATCT GCAGCACTTA CAGGACCATG TACCGCCTTA TGTGGCGGGA GGTGAGGCG	60(
	GAGGTTCAGC AGACCCATGC AGTGTGCTCC CAGGGCTGGA AGAAGCGGCA CCCGGGGGCC	66(
35	CTCACCTCTG AAGCCATCTG CGCCAAGCCT TGCCTGAACG GAGGCGTCTG CGTTAGGCCT	720
	GACCAGTGCG ACTGCGCCCC CGGCTGGGGA GCCAAGCACT GTCATGTGGA CGTGGATGAA	780
40	TGTAGGACCA GIATCACCCT CTGCTCGCAC CATTGTTTTA ATACGGCARG CAGCTTCAMI	840
	TGCGGCTGCC CCATGACCTA GTGCTAGGCG TGGACGGGCG CACCTGCATG GAGGGGTCC	90(
	CAGAGCCCCC AACCAGTGCC AGCATACTCA GCGTVGCCST TCGGGARGCG GAAAAAGATV	96(
45	ACGCGCTCTG AAGCAGGAGA TTCACGAGCT GCGAGGCCCT TGAAGCGGCT GGAGCAGTG	1020
	NCCGGTCAGC TAGGCCCTAG NTCAGACAGT GCTGCCCGTG CCGCCTGAAG WGCTGCAGC	1080
		1140
50	AGAACAGGTG GITGAGCTGT GGGGCCGGGG TGACCGGATC GAATCTCTCA GCGACCAGGI	
	GCTGCTGCTG GAGGAGAGGC TAGGTGCCTG CTCCTGTGAG GACAACAGCC TGGGCCTCGG	1200
55	CGTCAATCAT CGATAAGAAG CCTCTACAGC ACCCCTGCCC CCTAATTTAT ACAGAAACCC	1260
55	GACCCACTAA ICCTCT95GA TTGGCCGACT GTGAGCTGCA GATAAGGCTA TCAGCCACCA	1320
	AAGAGCAATG AACAATGSAA ACTICAGAGA GCIGAAGAAA GGGGGAGGCC TGTGTTCTIK	1380

60 GCCTGCCCCT GAGTCTTCTG GCTGGGGGGA GGTTGGCTGG GCAAGAACTG CTTCTTCAAT 1440

	TCCTTAACAA	. ATGCAACCAC	CAACACCCAG	Actorororor	CICITIATTI	TOAGTTOTT.	150
5	TOCTOTIATO	CAGATAATTA	ACCALALACCA	AAACEDAA	ACTGGGTCCC	ACCOINTOCT	156
	TIT GOTOCOA	GCCIACCTCC	CCAGTIGIGG	GU.CAGGICI	GGAGTGAGAG	GIAGGNAGTY-	162
	GCTAATGCCN	CCAGGAAGAA	ATGAAAACTG	GAGAGAG TON	GGGGAAGCCT	CAACAGAAAA	168
10	AGAAATAAAT	TAAAAGCCCT	CCTATCCCCT	CCAGCCAGGG	TTOGTTCCTT	YTOAACOCOOT	174
	CCCAGGGGGC	AGAAGTGAGT	GCAGCAGCTG	AS PROTECTE	CTTCCCCTTG	TGTCTGGTGA	180
15	GATGGTYGCAG	CAGGGCTGCA	GGGGGCTGGG	TGEGGTCATG	TCCACTGAAG	AACTGTACTA	186
	TGGGGACAGA	AAACCAGAAA	TGTGGAGACI	GAACIGGTAT	CCCAGAGAGT	GIAGGACCCI	192
	GGGCATCTGG	GCAAGGGCAG	GCATGAGACC	TCTGAATTAG	AAGGGTCCAG	COCCCACTGA	1980
20	CAGGAGGCTA	CACTGGGAGG	DAAG TODAAD	CI GETGAGGA	AAGCTCCCAT	GATGAGCCTS	2040
	GGAGTGCTTC	AGGTATCAGC	DADCOADOTT	AG 990GAGAA	GTCCTCCTCA	CAAATGGATG	2100
25	AGTCCATTGA	ATCCATGGAC	TTTGGAGTGG	CGSSGATTTG	TTCCAAAGAA	TGGATGAGTC	2160
	CACTGGCCAA	TGTGGGGTAG	DADA TEDEBA	AAGACCACAT	AGGAAGAGAC	TOCACT GGG(-	2220
	ATGGAATGTT	CCCCTCCCTT	OTOGĐATOTO	AGTCACTGGA	GAT/GA/GGGGG	AGGCAACTGT	2280
30	CCCACAGACA	AFACAGTAGG	AGGTGGGGGT	CAAGAGTGGA	GACTGCACCG	AGGCAAGAGT	2340
	CCAMGGATGG	GGCCAAGAGG	GGGCAGJAGT	CGCCCCTGTAT	CCACATTICA	CTTCAGAAGT	2400
35	TGAAGATTCC	AAAGAGGAGA	ATAAGTGGGG	AGAGGGGAGA	CAAGGAAGAG	GGTTTFGCC	2.460
	TGCTTCAG3G	CCCACTGGGT	GGGTAGJTGT	GG YGAGGAAG	ATNGGGGACAG	A'I\GGGAGGA'-	2520
	AGCTCAGA 3C	CAGGGTTCAC	CCACCGCCCC	CALABOTTCTT	CA SATAGTCA	CCOCCOCCOC	:580
40	GGCCATCAGT	GGAGATTTCC	CGGAAAACAG	TGAAGCATGG	AGTGCCGGAC	TOTGTCAGCC	2640
	AGAGTTGGGA	CGTCATCTGG	TGTCAGCCT	TICCETYSGGCA	CTGGGGCAG	CACCGGCAC:	2700
4 5	TGACATTISTC	COGAGGTGAA	GOGACGOTOO	THUTTIGCAGE	AGAAGTCTTG	CTAGGAGGATO	2760
						GOCACGTTT -	1820
						GACCYTCACH	1880
50	ASTEROTTOT	GCAGGGTCTG	G90GACCTCA	GA: ATGACAC	TGAT/GGCATC	O 10037130G	1940
	Trocestosa	TGAAGATGAC	тоствослес	CARTAGGICA	GICCGCAGAG	CCAGCCCACA	3000
55	DESPOYMATE	TYPGCAATGG	GDACACAGOG	GOTTYG/CAGT	ACCTCCATCA	TOCCAAGCAG	3060
						GGAAGT GGTYG	3120
. 0						GCATTAGACC	3180
50	CAAGATCTTC	ATGTTCTCGA	CGTTGCGTCC	TYTHUACGGCA	CACACAGGGA	TGGCGA:3CA*	3240

	AGCCAGGAAG AGGATCCAGC CATTOTAGAA GGCCATCTTG AAGAAGTACT TGGCACTGGG	330
5	(-	330
10	(2) INFORMATION FOR SEQ ID NO: 67: (i) SEQUENCE CHARACTERISTICS:	
15	(A) LENGTH: 1535 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPCLOGY: linear	
	(xi) SEÇUENCE LESCRIPTION: SEQ ID NO: 67:	
20	GBCACGAGGT CAAGCGAAAG GATTICAAGG AACAGATCAT CCACCATGTG TICACCATCA	6
	TTOTCATCAG CTTTTCCTGG TTTGCCAATT ACATCCGAGC T9933ACTCTA ATCATCGCTC	12
	TGIATGGACT CTTCCGATTA CCTGCTGGAG TCAGCCAAGA TGTTTAACTA CGCGGGAATGG	18
25	AAGAACACCT GCAACAACAT CITCATCGTC TTCGCCATTG TTTTTATCAT CACCCGACTG	24
	GTCATCCTGC CCTTCTGGAT CCTGCATTGC ACCCTGTGT ACCCACTGA GCTCTATCCT	30
30	GOOTTOTTTG GCTATTACTT CTTCAATTCC ATGATGGGAG TTCTACAGOT GOTGCATATC	36
30	THORSEGOOT ACCTUATITY GOSCATISCO CACAASTICA TAACHISAAA GOTISTAGAA	42
	GATGAACBCA GTACCGGGAA GAAACAGAGA GCTCAGAGGG GGAGGAGGAC TOEGCACGCCCACACACACACACACACACACACACACACACA	48
35	GAPBAGCAAA GAGCCGGCC CTAGCCAATG GCCACCCAT CCTCAATAAC AACCATCGTA	5 4
	AGAATGACTG AACCATTATT CCAGGTGGCC TOCGTGAGAACTA GCAAAAAAC CGAAGAACTA	60
	POADASTEAA ASASOAAAAA DODOTOTOGA ATTTOACTEG DATATODOOT OOCTOSOOOD	66
40	AGAGTTOTOT GCATOCTOCO TOOTTOGTTO TOACCOAGTT GCCTTTAAAC CAAATTOTAA	721
	CCAGCCTATC CCCAGGTAGG GEGACCTTGG TTATATTCTG TTAGAGGEGG ACGETCGTAT	78
45	TTTCCTCCCT ACCCCCAAG TCATCCTTTC TACTGCTTTT GAGGCCCTCC CTCAGCTCTC	84
	TGTGGGTAGG GGTTACAATT CACATTCCTT ATTCTGAGAA TTTGGGCCCCA GCTGTTTGCC	900
	TITHGACTOOC TGACCTCCAG AGCCAGGGIT GTGCCTTATT GTCCCATCTG TAGGGCTCAT	961
50	TCTGCCAAAG CTGGACCAAG GCTAAGCTTT CTAAGCTCC TAACTTGGGG CAGAAACCAA	102
	AGCTGAGCTT TTAACTYTCT CCCTCTATGA CACAAATGAA TTGA PGSTAG GAGGAGGGTG	108
55	CACATAACCC TTACCCTACC TOTGCCAAAA AGTGGGGGCT GTACTGGGGA CTGCTCGGAT	
رد		
	GATETTTETT AGTGETACTT CTTTECAGETC TECCTETAGC GACAGETETA AGATETGACT	120
60	GEOTOGTCCT TICTCIGGCC MATTACOCACT TOCCICTTCT CTAGAGGTAG GATAGGTGGT	1260

	TTGGAGLAGA AUGGUAACIA AULANIT TUANTIANIA AALAUTTGGG GITTTGGTTT	1320
	TANAGOCAGA ATTACCHUTTA GCACCITAGCA TTTCAGCAGA GUGACCATT TAGACCAAAA	1380
5	AAAAITA ATAAAAAIA AATTAGAAAA TYAAAAITTO CYYYYEGGOA ATTGTGATGT	1440
	CATGGCAATA AGTGTTAGAC TATTAGGAAT TGAGAAGGGG GATCAACTAA ATAAACGAAG	1500
10	AGAGICITTIC TIATGUAAAA AAAAAAAAA AAAAA.	1539
15	(2) INFORMATION FOR SEÇ II NO: 68:	
20	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1244 base pairs (E) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear 	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:	
0.6	GGGCACCCAC CAGCGGRGCT GARTTCAGCG CGCACCTATG GGCTRGCTAC CAGGACATGC	6
25	ACCORCAGE SCARGACCTO CIGNOCOCE AGETOTECAS TOTOCTEAAC COAGCAGEAC	12
	TOTACGCCAA CAACGAGATC AGUUTGCGTG ACGTTGAGGT CTACGGCTTT GACTACGACT	2.8
30	ACACCCTGGC CCAGTATUCA GACYCACTGC ACCCGAGAT CTTCACTACC GCCCGTGACA	24
	TOCTGATOGA GCACTAGAAG TAGOCAGAAG GGATTCGGAA GTATGACTAC AACCCCAGCT	30
35	COORDADYTA RAARTOT TOORAGAARA OTTADAATTA CAUCTORROT ROOTAGUTT	36
5.	TOCACTACIST GCAGOTTGGGS ACAGOTTACA GGGGCCTCCA GCCTGTGCCA GACGAGGAGGAGG	42
	PRETERED TORRESTATE ACCEPTAGE ACCEPTAGE TORRESTATE TORR	48
40	GCAAGEGTCC CTCCATTAAG CAGTICATEG ACATCTTCTC GCTACCEEAG ATGGCTCTGC	54
	TETCCTGTSI GST/SEACTAC TTT/CIGGGCC ACAGCCTGGA GTT/SEACCAA GCACATCTCT	60
45	ACAAGGACGT GACGGACGCC ATCCGAGACG TGCATGTGAA GGGCCTCATG TACCAGTGGA	66
7.	TOGARDAGGA CATREAGARS INCATOCINGA GAGGEGATGA GACGITITECT GTCCTGAGCC	72
	GOCTOPPERGO COLATORADA CAROTTOCO TOATOACOAA CASTOCRITO ARCTOCRIAG	78
50	TETTACHEST DIASOULUID ACOROSSITA ENCONOSERDO ENACACERDO IASDECANA	84
	ADTABOTICAA AACAOMITO DAACDOGGOO AGTICATITOT TOGACOCAAA AACTGOACO	90
55	GAAGGGCTCA CTYCAGTGGG ACCUGATCAC COGCTYGGAA AAGGGCAAGA TCTATCGGCA	96
٦,	GEGAAACOTG TITGACTICI TACGCTIEAC GGAATGEOET GECCCOCGCG TECTETACTI	102
	COORDACCAC CTCTATAGTS ATCTGCCOGA TCTCATGCTG CGGCACGGCT GOCGCACAGG	108
60	CGCCATCATC CCCGAGCTGG AGCGTGAGAT CCGCATCATC AACACGGAGC AGTACATGCA	114

	CTCGCTKACG TGGCAGCAGG CGCTCACGGG GCTACTAGAG CGCATKCAGA CCTATCAGGA	1200
5	CGCGGAGTIG AGGCAGFICT TGCTTCTTIG ATGAAAGANC GNNI	1244
10	(2) INFORMATION FOR SEC ID NO: 69: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1292 base pairs (B) TYPE: nucleic acid	
15	(C) STRANDEDNESS: double (I) TOPOLOGY: linea:	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:	
20	GGCACGAGCA GCGACGCBAC TCTGGTGCGG GCCGTCTTCT TCCCCCCGAG CTGGGCGTGC	€(
20	GCGGCCGCAA TGAACTGGGA GCTGCTGCTG TGCCTGCTGG TGCTGCTGCTGCTG	120
	CTCTTGGTGC AGCTGGTGGG CTTCCTGAGG GCTGACGGCG ACTGGAGGCT ACTATGGGCC	180
25	GAGTGGCAGG GACGACGACCACAGAATA GAGCGCTGACTA TYPETGGTGACGACAGAA	24(
	GCCTCGAGTG GAATTGGTGA GGAGCTGGCT TACCAGTTGT CTAAACTAGG AGTTTCTCTT	300
	CTCCTCTCAG CCAGAAGACI GCATGAGCTG GAAAGGGTGA AAAGAAGATG CCTAGAGAAT	361
30	GGCAATTTAA AAGAAAAAAA TATACTTGTT TTGCCCCTTG ACCTGACCGA CACTGGTTCC	420
	CATGAAGCEG CTACCAAAGC TETTETEEAG GAGTTTGGTA GAATCGACAT TETGGTCAAC	480
35	AATGGTGGAA TGTCCCAGCG TTCTCTGTGC ATGGATACCA GCTTGGATGT CTACAGAAAG	541
	CTAATAGAGC TTAACTACTT AGREGORGE TCCTTGACAA AATGTGTTCT GCCTCACATG	£ ()-
	ATCGAGAGGA AGCAAGGAAA GATTGTTACT GTGAATAGCA TOUTGGGTAT CATATCTGTA	€ 6.1
40	CCTCTTTCCA TTGGATACTG TECTAGCAAG CATGCTCTCC GEGGTTTTTT TAATGGCCTT	720
	CGAACAGAAC TTGCCACATA COCCAGATATA ATAGTTTCTA ACATTTGCCC AGGACCTGTG	78(
45	CAATCAAATA TTGTGGAGAA TTGGGTAGGT GGAGAAGTCA CAAAGACTAT AGGCAATAAT	840
	GGAGACCAGT CCCACAAGAT GACAACCAGT CGTTGTGTGC GJCTJATJTT AATCAGCATG	900
	GCCAATGATT TGAAAGAAST TTSSAICICA GAACAACCTT TCTTCTTTAG TAACATATTI	960
50	TADDAAAAADA ADEDETAEJA CAACCAATAD DISEICCEED TOCAACCATAC ATAACDOTD	1020
	TGAGAACTTT AAGAGTGGTG IGGATGCAGA CTCTTCTTAT TITAAAATCT TTAAGACAAA	1080
55	ACATGACTGA AAASAGCACC IGTACTTTTC AAGCCACTGG AGSGAGAAAT GGAAAACATG	114(
	AAAACAGCAA TCTTCTTAINE CTTTTEGAATA ATCAAAGACI AATTTEGGAT TTTACTTTTT	1200
60	AATAGATATG ACTINGCITC CAACATGGAA TGAAATAAA AATAAATAAT AAAAGATTGC	126(

يقاضيه والمارات

216

	CATGAATOTT GCAAAAAAA AAAAAAAAAA AA	1251
5		
	(2) INFORMATION FOR SEC ID NO: 70:	
10	(i) SETVENCE CHARACTERISTICS: (A, LENGTH: 1031 base pair: (E) TYPE: nucleic acid (C) STRANTEENESS: double (D) TOPOLOGY: linear	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70.	
1.	GGGCTGTTGC TTTTGAACAG AACCCTATAL TACTCTCTG GGATCTGAGT TFCTGCAGG.	€ (
	PROTECTION ADDITION AND ADDITIONAL PROPERTIES OF A PROTECTION	11(
20	GCALATTUTA AGTIGCOATA TIMAACATCA TOCOCACTOGO AGTIGTITAT TITTATACAC	18(
	AUGITGGCIG GCTTCAGTTI TTP3CTGIAGC CCIAGAGCAC TTTGTTTGIG GGAGGCTGG	24(
25	CTCTTGCCIA CCICCTTGCA TGSACAGGGG GATSAATATT TACTFTCCCA CCTCCTTGCT	300
23	TTTTCTTTCA CTSATACCAC TGAATGGAAC TGSTGCTGTG ACTCCTGCTG CTGGGGATTT	360
	ATMSTCCCOGAG ACCIMIAGODI GEORGAGIES AGODIGAGAC CIGGACAACA GOTICATEGET	41(
30	ATGCATGARA GAGAGAGTG COACACAGO AGAGGAGAACA GIAAGAGAC AGGGGGACATT	48)
	ATTITOGGAA AGSCTSTCCC GGGCTGTTAC TGTCTCTTCT GGTTATAAAG CAGACATGTG	540
35	GCCATCTTTT CCGCAGGTTA GAGTGGGCTC CTTPCTTTTT: GGAATCCTTT TCTTCTCCTT	60)
35	TVRSTAGGAGO TOCCTURECTO CARGESTICO GECACCAGOS TVCTOTECTET GITTSCGCAGI	660
	GENGTQBOOK GENAGERET TETTYCTGOO TWOOTYAAAG ARAGYGETOT GERGATYGAG	7.0
40	ATGAGAAACA ALAGGETETE CITECAGACAA INAGGEATTE TETECTECTE CITECCATTEI	78(
	CONTRACTO TRADERIADAD DEFENACIONA ERATERIORA BACCEANACT DADOTOTOTACO	841
4.5	CTACTOTTAG SITIVICISI GIGATCOTTI CCCTCCCTSI CGCCCACTCC TCCCTCCTCI	90
45	MAGTETGAE DETETGEAT GEDACIATIT ACTACAGE TETGTGTGC GACTETGEB	9 5(
	GCATTLAGTI CASAGISTAN GGYCTTTGGS CIGAAATAAA AIGCAACTAT TIAAAAAAA.	102.
50	ALLALAAA L	113.
55	(2) INFORMATION FOR SEC ID NC: 71:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 055 base pair: (B) TYPE: nucleic acid	

(C) STRANDEDNESS: double

(D) TOPCLOGY: linear

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:	
5	AGCTATTGAC ACTTCCTTGE TEAGOCEAGT GAGCOCTCCCT TGGCGCT TGGCGCT TGGCGCTCAG	61
	GCSGCSACCA TEGCSTATCA CESCCTCACT GTGCCTCTCA TESTGATSAG CSTGTTCTGE	120
10	GGCTTCGTCG GCTTCTTNGGT GCCTTGGTTC ATCCCTAAGG GTCCTAACCG GGGAGTTATC	180
	ATTACCATGT TGGTGACCTG TICAGTTTGC TGCTATCTCT TTTGGCTGAT TGCAATTCTC	24
	GCCCAACTCA ACCCTCTCTT TGGAACCGCAA TYGAAAAATG AAACCATCTG GTATCTGAAG	300
15	TATCATTGGC CITGAGGAAG AAGACATGCT CTACAGTGCT CAGTCTTTGA GGTCACGAGA	3€(
	AGAGAATGCC TTITAGATGC AAAATCACCT CCAAACCAGA CCACTTTTCT TGACTTGCCI	420
20	GTTTTGGCCA TIAGCTGCCT TAAACGTTAA CAGCACATTT GAATGCCTTA TTCTACAATG	480
20	CAGCGTGTTT TCCTTTGCCT TTTTTGCACT TTGGTGAATT ACGTGCCTCC ATAACCTGAA	540
	CTGTGCCGAC TCCACAAAAC GATTATGTAC TCTTYCTGAGA TAGAAGATGC TGTTCTTCTG	600
25	AGAGATACGT TACTCTCTCC TIGGAATCTG TG3ATTTGAA GATGGCTCCT GCCTTCTCAC	660
	GTGGGAATCA GTGAAGTGTT TAGAAACTGC TGCAAGACAA ACAAGACTCC AGTGGGGTG	720
30	TCAGTAGGAG AGCACGTTCA GAGGGAAGAG CCATCTCAAC AGAATCGCAC CAAACTATAC	780
30	TTTCAGGATG AATTTCTTCT TTCTGCCATC TTTTGGAATA AATATTTTCC TCCTTTCTAW	840
	REAAAAAAA ANANI	23
35		
	(2) INFORMATION FOR SEQ ID NO: 72:	
40	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 1274 base pairs (B) TYPE: nucleic ació	
	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:	
	CCCAGACTT AGACTCTGGA AAAGGCAACC AGCTTGCCCG TAACTGCCTG CTGGAATGCC	6
50	TGTGCCTCCA CACGGGTCTG GGCATCCGGA CTGATAACCA GCCGGCCAGA CTGAGGGATG	12
	POTOTTTO AGATAACK DECOMMEND DECACTOR DECACTOR AGAAATGAG CTTTCTGTG	18
	TOTOTIGGAC TOTISGOTISCO TOTISGOCTO POTISTOTOTO TOTTTOTIGG TOTOTOCOTI	24
55	TOTOCTOCTO AGGOTEGICT TTOTOTTIGG TECACACTIA GITATIGITG IGAGCAAIG	30
	AAGTTCAAAG GAACTCCCTC TCCAGCTCTT CTGAATCTTG GGACACAGCC TAAAAAGGA	36
60	PARABOTTE FIGURACIAT FORESCENCE CONTRACTOR TANGERS ARTRACTAN	42

	CONTROLATO AADALDAARA AAAADAAROO DAADDINTTO TYDDINTTO- DEEVEDTARD	48(
5	THE AGAGE CONTINUE TO THE AGAIN THE CONTINUE OF THE AGAGE THE AGAGE TO THE AGAGE THE A	54(
2,	TOACCCUA OTCOCCTAG GAGAAAAUTO TGAGTYTCEG GAUATCCAAA DATETGAAAQ	ϵ cc
	WESTCADADE OSTOTADADA DARREGERADA ERRITORREGER TITREREGEROTOA DRAMADEGERAT	660
10	GATATTICTO (CASATTACA GTYTCTIGOS GCCIAAACAG GTTAGGTAGA CTATAGGCTY	720
	TITPEETTEE ETTTAGEATT ETEPACITOTI ETAGOAT EGITITCOAN OUTGOAGEN	780
15	THEIGHTS THEFTITE COMMISSED THEFTHETH DOUGHTTA AASAAAAGO.	840
15	AMAGGC:GC: GCGAGTCTH-GNAGAGGCI CTCLANGGAI GNAGCATATC GAAGATAATT	900
	AMPRACEGAD TODITOTODOS TAGRICISAS TOTITITANTA SEMATITITA DOTDATATET	960
20	GAGGGGTTON AGREGAAAGNO ACNOACTTIN AGREGAGTTO CTNOCCAGAGE GAGGGGAN	1020
	GROATGREAT HINGAGGIFF GORACACACT CIRTRICINIT TOAARGCIGG GOGREFFER	1080
25	CUTCCAGAG: CTCTCTG993T CTCAGATGTC CATCTGCCAC CTCTTGTTAA GGCTCTAGCC	1140
دع	AGAAGGSAG: CTSAGGGTAG AAGAAGTTA TTCCCGAAGA AAAAAAGAAT GAAAAGTCAT	1200
	TOTACTGAAA AAAAAA TOTTTAAAAAS TUACTATTAA AAAGTAAAAA AAAAGG KAK-	12:60
30	COCGGTACC AATT	1274
35	(2) INFORMATION FOR SEQ ID NO: 75.	
40	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 688 base pair: (E) TYPE: nucleic ació (C) STRANNEDNESS: double (E) TOPOLOGY: linea:	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:	
45	GBCACGAGTS SAGSCAATGC CASCTCCAGG ACAGAGGCTC AGSTGCCCAA CGGGGCAAGG	60
	AGRICCAGGG: CITGTGTCTG TIVAAGTCAG GCTTCCCC 36 COCTCGCGCA CAGCGCTTC	120
50	CONTROTADO STORRETORS DESEMBARAT DADROADOND ADODERSON CORACEROMA	180
50	GACCTTCCT: CURRINGING TCACCCTRRC CACSTCTHCA CACRRCTGCA CARAAACTT	240
	CONTORTO ACADEACOR ACARROPACA COCREGORO ATOTACCAS EDERCEDADO	300
55	GTGCTGAAGO (+3AGGCTGCT GCAGCCCTCG CGCCGGGTCA AGCGCTCGCG CCGGAGACCC	360
	APPENDAGO APTRACONO DAGAGOGRAA FENDORAGA AARRONANA DAGAGORA ORGANIZATO	420

COTECCACTS TESCOTORES CTCCTCCCES CECCGCEAGG COGCEACCTC TESCACGTE 480

	ACCESSORSE TAUSSETTER COSTROSES ROSCESCAL TR	GCCGAGCA CIGCGGGGGC 540
	TIMICCTOCTT GTTGGTTGCT GAGTUGGCGG CCAAGGGGAG AAA	AAGGAGCC GCTTCTGCCT 600
5	KEA TDATYTTTAI IAAATIAAID TTTDOOTOAA AACOOTTOOO	AAAAAAA AAAAAAAA
	MARAKA AAAAAAA AAAAAAAA	688
10		
10		
	(2) INFORMATION FOR SEQ ID NO: 74:	
15	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1890 base pairs (B) TYPE: nucleic acid (C) STRANDEDMESS: double (D) TOPOLOGY: linear	
20	(xi) SEÇUENCE DESCFJFTJCN: SEÇ ID NO: 74	4 :
	GAGCAGSAGA GAAGGCACCCC CCCCACCACCCC CCTCCAAAGC TAA	ACCCTTGG GCTTGAGGGG 60
25	AAGAGGETGA CTGTACGTTC CTTCTACTCT GGCACCACTC TEC	CAGGCTGC CATGGGGCCC 120
	AGRACCOCTO TOOTCATCTT GTTCTTTTTG TCATGGTCGG GAC	CCCCTCCA AGGACAGCAG 180
	CACCACCTTG TGGAGTACAT GGAACGCCGA CTAGCTGCTT TAG	BAGGAACG GCTGGCCCAG 240
30	TODCAGGACO AGAGTAGTOG GCATOTTGCT GAGTTGCGGG ACT	TTCAAGAA CAAGATGCTG 30C
	CCACTGCTGG AGGTGGCAGA GAAGSAGCGG GAGGCACTCA GAA	ACTGAGGC CGACACCATC 360
35	TCCGGGAGAG TGGATCGTCT GGAGGGGAG GTAGACTATC TGG	SAGACCCA GAACCCAGCT 42(
5.	CTSCCCTGTS TAGASTTTGA TGAGAAGSTG ACTSGAGGCC CTG	993ACCAA AGGCAAGGGA 480
	AGAAGGAATG AGAAGTACGA TATGGTGACA GACTGTGGCT ACA	ACAATOTO TOAAGTGAGA 540
40	TCAATGAAGA TTCTGAAGCG ATTTSSTSGC CCAGCTGGTC TAT	rgaaccaa gaatccactg 600
	GUULAAACAG AGAAGATCTA COTOTTAGAT GUUACACAGA ATO	SACACAGC CTTTSTCTTC 660
45	CCAAGGCTGC GTGACTTCAC CCTTGCCATG GCTGCCCGGA AAG	SCTTCCCG ASTCCGGGGTG 720
, .	COUTTOCCCT GGGTAGGCAC AGGGCAGCTG GTATATGGTG GCT	TTTTTTA TTTTGCTCGG 780
	AGRICTICTE GAAGACCTGG TGGAGSTRGT GAGATGGAGA ACA	ACTITION ASSERTED ASSETTED
50	TTOCACCTGG CAAACCGAAC AGTGGTGGAC AGTTCAGTAT TOO	CCAGCAGA GGGGCTGATC 900
	CCCCCTACG GCTTGACAGC AGACACCTAC ATCGACCTGG CAG	GCTGATGA GGAAGGTCTT 960
55	TGGGCTGTCT ATGCCACCG GGAGGATAAC AGGCACTTGT GTG	CTGGCCAA GTTAGATCCA 1020
55	CAGACACTGG ACACAGAGCA GCAGTGGGAC ACACCATGTC CCA	GAGAGAA TGCTGAGGCT 1080
	GOOTTIGICA TOTGTGGGAC COTOTATSIC GTGTATAACA COO	COTOCTGC CASTUGGGCC 1140
60	CECATOCAGT COTOCTTGA TECCACCOGA COOTGACCCO TEA	ACGOGCA GCACTCCCTT 1200

	PACOCOGA GACOCOCATA ATOGOCOTOGO ACCORTACIONO TESTATADAS SOCCOSTITUTA	1260						
4	TOTATHSCOTS GSATGATGSC TATOAGGATTG TOTATAAGCT GSAGATGAGG AAGAAAGAGG	1320						
•	MATAUTTAD ATACOCCOAC TOTTTC DACO CTUTTCOUT CACCOACA CAUTERARRA	138(
	MAAADYDEEST EVTEADORRE TETAAADYDD YYD YYDYYTAAAY LADODOGADA	144						
10	CONCIATATI TYMAGOCAAT +49.AATCAAA TYCTYTCOTO TYCTYTTOTO CATAGOCATA.	1500						
	TOCAGATOTT GAGTAATOCT TT:AGAGACOTO GAAGAGTGA AACCTODAAT GTTCCCCCC	156						
15	AATOTOCTOD COCATOTOAA CAACTTOO GOTAAGGATG COCCAGACCO AGGGTTCTAA	162						
10	PTTOAGTEAD AGTOCOUTTO TTETBAGEAG GEAGGAGHAN COCGODAGGAGADAG	1680						
	ASTECCTES TOASTOUTES TOTOSTOTE SEASOAGEA ASEASOAGEA AGASSEAGED	174						
20	CTGTYCTYAG GAACAGACT TTTTCTAT TGTTTTTGTAI TYCAACATTT TGCATTAAAA	180						
	GGAAAATYCA CTGCAAAAAA AAAAAAAAA AAAAAAAA AAAAAAACG CACGAGGGGK	18€						
25	GCTCCCGTAC CCDATHGCCC TCACATYCAT	189						
2."								
	(() ANTIONNETON FOR OTO 3TO NO. 76.							
30	(2) INFORMATION FOR SEQ ID NO: 75:							
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1133 base pairs (E) TYPE: nucleic acid							
	(E) TYPE: nucleic acid (C) STRANIEDNESS: double							
35	(E) TOPCLOGY: linear							
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:							
40	GOOGGTOTGA GTGCAGAGOT GOTGTCATGG CGGCCGCTCT GTG33GCCTCT TTTCCCCGTCC	6						
	TECTGOTGOT GOTECTATOS GOESEATOTOS AGAGOTOGGA GETECOGGS GOTGOTGOTG	12						
	AGRICATICGGG AGRICAGING CICCGCATAG GAGATCGCTT CAARATTGAG GGGCGTGCAG	18						
45	TTSTTCCAGG GETSAAGCCT CARSACTGGA TCTCGGCGGC CRGASTRETG GTAGACGGAG	24						
	AAGAGCACGT CGGTTTCCTT AAGACAGATG GGAGTTTTGT GGTTCATGAT ATACCTTCTG	30						
50	GATCTTATGT AGTGGAASTT GTACCTCCAG CTTACAGATT TSATCCCSTT CGAGTGGATA	36						
50	TCACTTOGAA AGSAAAAATS AGAGCAAGAT AISTGAATTA CATCAAAACA TCASAGGTTS	42						
	TCAGACTGCC CTATCCTCTC CAAATGAAAT CTTCAGGTCC ACCTTCTTAC TTTATTAAAA	48						
55	GRIGARIORIG GRICTEGRACA GARTITUTAN TRANCOCANT GRITATRATG ATGITTOTT	54						
	CTTTATTGAT ATTTGTGCTT CTGCCTAAAG TGGTCAACAC AAGTGATCCT GACATGAGAC	60						

GGGAAATGGA GCAGTCAATG AATATGCTGA ATTCCAACCA TGAGTTGCCT GATGTTTCTG

	AGTTCATGAC AAGACTCTTC TCTTCAAAAT CATCTGGCAA ATCTAGCAGC GGCAGCAGTA	72(
	AAACAGGCAA AAGTGGGGCT GGCAAAAGGA GGTAGTCAGG CCGTCCAGAG CTGGCATTTG	780
5	CACAAACACG GCAACACTGG GTGGCATCCA AGTCTTGGAA AACCGTGTGA AGCAACTACT	841
	ATABACTTGA GICATCOOGA OSTIGATOTO TIACAACTGT GIATGITAAC TITTIAGCAC	904
	ATGTTTTTTT OTTGGTACAC GAGAMAACC AGCTTTCATC TTTTCTCTGT ATGAGGTCAA	96.
10	PATTATTAA TAATTAADATATO OTETBADATAA TAATTAATTAA	102(
	AACTACTATA CATTATGTAT ATTAATTAAA ACATCTTAAT CCAGAAAAA AAAAAARAA	1080
15	AACTCGAGGG GGGGCCCGGT ACCCAATITN CCAAATGGGA GTCGTAAAAA ATC	1133
20 25	(2) INFORMATION FOR SEQ ID NO: 76: (i) SEQUENCE CHARACTEFISTICS: (A) LENGTH: 585 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
20	(xi) SEQUENCE LESCRIPTION: SEQ ID NO: 76	€(
30	AIGTTTACAA TGTTGIGIAT AAATGGGACA ACTCCTCGCC CTCTACCIGT CCCCTCCCCC TTTGGTTGTA TGATTTTCTT CTTTTTTAAG AACCCCTGGA AGCAGCGCCT CCTTCAGGGT	120
	TOGCTGGGAG CTCGGCCCAT CCACCTCTTG GGGTACCTGC CTCTCTCTCT CCTGTGGTGI	180
35	CCCTTCCCTC TCCCATGIGC TCGGTGTTCA GTGGTGTATA TTTCTTCTCC CAGACATGGG	240
	GCACACGCCC CAAGGGACAT GATCCTCTCC TTAGTCTTAG CTCATGGGGC TCTTTATAAG	300
40	GAGTTGGGG GTAGAGGCAG GAAATGGGAA CCGAGCTGAA GCAGAGGCTG AGTTAGGGGG	360
40	CTAGAGGACA GTGCTCCTGG CCACCCAGCC TCTGCTGAGA ACCATTCCTG GGATTAGAGC	42.0
	IGCCTTTCCC AGGGAAAAAG TSTCGTCTCC CCGACCCTCC CGTGGGCCCT GTGGTGAT	480
45	GCTGTGTCTG TATATTCTAT ACAAAGGTAC TTGTCCTTTC CCTTTGTAAA CTACATTTGA	54(
	CATGGATTAA ACCAGTATAA ACAGTTAAAA AAAAAAAAA AAAAA	585
50	CETGGATTAA ACCAG,ATAA ACAGITAAAA AAAAAAAAA AAAAA	
	(2) INFORMATION FCR SEQ ID NO: 77:	
55 60	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 577 base pair: (B) TYFE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	

	OXIC SEQUENCE DESCRIPTION: SEQ ID NO: 70:	
	OBSCHOLOGICANTE CORRECTION TO SETTING CALING THE SECURITY OF THE CORRESPONDED	€-
5	GSTGCCTCAG CTTCCTTCTS ATMOGRACHT TCCTGT.AGT TTCCCAGACA GTCCTGGCCC	12:
	ORACTIONA ISTINTINA AUTHORIDAA UNGOANNIT UTBOTINATOA DEMAGGITSEA	18
10	COCASCACIO ECCEACIDATE STECACACA GACTACIDE STECACACACACACACACACACACACACACACACACACACA	24
10	COLOCATO TOTESTORE TAGGERA ESSTENDO A TRADOTODO	3 (0)
	MEATEACOA CENTETERIO SETMACACCO EMADINARMA COEACAROID ETAGOTAROO	36
15	COOFFERTY SEFERTSWAY PATATEFICA TOATUAKEDD SADSAAAAST SOEMSBUESOS	42
	AGRIGITORES TOTTOLOGIC TOTTOCOC TOTTOCOCACTOCIC CIGACOTTOC	48
20	YTTAIAAIAA TT: GOTAAAA IST TIDOODT TIGTTODSAG IDIDTYTOJAA AITYTIDYNYD	54
20	MACKISICAA CAACAALAMA NAAAMAWAAA MACTIGA	57
25	(2) INFORMATION FOR SEQ ID NO: 78:	
	(i) SEQUENCE CHARACTERISTICS:	
20	(A) LENGTH: 2278 base pairs	
30	(E) TYPE: nucleic ació (C) STKANDEDNESS: doubl∈	
	(D) TOPOLOGY: linear	
35	(xi) CEQUENCE DESCRIPTION: SEC ID NO: 78:	
٥.	GTAATTOGGC ACGAGGCGC CAACATGGCG GGTGGCCCCCC GCGCCCCACA SCTAACGGCG	f
	CUPPUTBARCO COTABARDOSC GGOTATGATO GOGACGATAG ANCONGARGA GACCGCGCT	11
40	- PERSONAL ACCEPTOR ACCEPTOR ACCEPTOR ACCEPTAGE OF ACCEPTAGE OF ACCEPTAGE AC	18
	GGIGASTYSA INSTYSAAATY TIACSCCCCA TWSTGTCCAT CCTGCCASCA GACTGATTCA	24
45	GAATGGBAGG CTTTTBIAAA GAATBGTBAA ATACTTBAGA TCABTGTBGG GAAGGTAGAT	30
٦.	GTUATTEAAG AACUAGSTIT GAGPSGCCGC TICTTTOTOA COACTOTOCOAGCATTTTT	3+
	CATGCAAAGG ACSSSATATT COGCOGTTAT CGTSSCSCAG GAATCTTCGA AGACCTSCAG	41
50	-DOCCTAAAD DICEDIDADE DICEDAECTE ACCASSERVA AACAASADE TOTATAAA	4 ⊱
	GOTTOTITAA CUATUTUGG AATGUOTGGT CTTTUTAGCA TOTTOTGGCAA GATATUGGCAT	£1
5.5	STYPTOTES ATTOTTOTES TIGETOSTIA APPTTENDAD TRADACCUTA TOAASAACC	60
55	GTOATAGCOA CCTTGGTTTT TGGGCTTTTT ATGGGTCTGG TCTTGGTGGT AATATCAGAA	66
	TGTTTCTATG TGCCACTICC AAGGCATTTA TCTGAGGGTT CTGAGGAGAA TGGGAGATCA	72
60	GAGGAGGOTO ATAGAGOTGA ACAGTTGONG GATGOGGAGG AGGAAAAAGA TGATTOAAAN	78

	GAAGAAGAAA	ACAAAGACAG	CCTTGTAGAT	GAT GAAGAAG	AGAAAGAAGA	TCTTGGCGAI	84
5	GAGGATGAAG	CABAGBAAGA	AGAGGAGGAG	GACAACTTGG	CIBCIGITGT	GGATGAGGAG	90
	AGAAGTGAG3	CCAATGATCA	ADDDDCCCCD	GGAGAGGAGG	GIGTGACCCG	GGAGGNAAGT	96
	AGAGCCTGAG	GARGETGAAR	AAGGCATCTC	TGAGCAACOC	PTPSIACOCECT	ACACAGAGGC	101
10	GSTGGAAGAC	TOUTTGAGGO	AGCSTAAAAG	TCAGCATGCT	GNOAAGGGAC	TGTAGATTTA	108
	ATGATGCGTT	TICAAGAATA	CACACCAAAA	CAATATGTCA	GETTCCCTTT	GGCCTGCAGI	114
15	TTGTACCAAA	TTTAATTT	TTCCTGAATG	AGCAAGCTTC	TOTTAAAAGA	TGCTCTCTAG	120
• •	TCATTTGGTC	TOATGGCAGT	AAGCCTCATG	TATACTAA93	AGAGTCTTCC	AGGTGTGACA	126
	ATCAGGATAT	AGAAAAACAA	ACGTAGTGTN	TGGGATCTGT	TYI BIJAGACTG	GGATGGGAA(132)
20	AAGTTCATTT	ACTTAGGGGT	CAGAGAGTCT	CGACCAGAGG	ABBOCATTCC	CAGTCCTAAT	138
	CAGCACCTIC	CAGAGACAAG	GCTGCAGGCC	CTGTGAAATG	AAAGCCAAGC	AGGAGCCTT(-	144
25	GNTCTGAGGC	DAAAOCCCTA	TGTAACGTAG	ASCOTTOCOA	TOCTTTTCTT	GTGTAAAGTA	1500
	TTTATTTTTG	TCAAATTGCA	GGAAACATCA	GGCACCACAG	TGCATGAAAA	ATCTTTCACA	156
	GCTAGAAATT	GAAAGGGCCT	TGGGTATAGA	GAGCAGCTCA	GAAGTCATCC	CAGCCCTCT:	1628
30	AATCTCCTGT	GCTATGTTTT	ATTTCTTACC	TTTTAATTTT	COAGCATTTC	CACCATGGG:	168
	ATTCAGGCTC	TOTACACEOT	TCACTATTAT	CTCTTGGTCA	GAGGACTCCA	ATAACAGCCA	174
35	GGTTTACATG	AACTGTGTTT	GTTCATTCTG	ACCTAAGGGG	TTTAGATAAT	CAGTAACCA"	180
	AACCCCTGAA	GCTGTGACTG	CCAAACATCT	CAAATGAAAT	GTTGTGGCCA	TCAGAGACTY:	186
	AAAAGGAAGT	AAGGATTTTA	CAAGACAGAT	AAAAAAAT	TTGTTTTGTC	CAAAATATAG	192:
10	TTGTTGTTGA	TTTTTTTTA	AGTTTTCTAA	GCAATATTTT	TCAAGCCAGA	AGTCCTCTAA.	1980
	GTCTTGCCAG	TACAAGGTAG	TCTTGTGAAG	AAAAGTTGAA	TACTGTTTTG	TTTTCATCTC	204
15	AAGGGGTTCC	CTGGGTCTTG	AACTACTTTA	ATAATAATA	AAAAACCACT	TCTGATTTTC	210
	CTTCAGTGAT	GNECTTTTGG	TGAAAGAATT	AATGAACTCC	AGTACCTGAA	AGTGAAAGAT	2160
	TIGAITTTGT	TTCCATCTIC	TGTAATCTTC	CAAAGAATTA	TATCTTTGTA	AATCTCTCAA.	2221
50	TACTCAATCT	ACTISTAAGTA	CCCAGGGAGG	CTAATTTCYT	AAAAAAT	AAAAAAA	2278

55 (2) INFORMATION FOR SEQ ID NO: 79:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1143 base pair:

(B) TYPE: nucleic acid

60 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

	(X1) SEQUENCE DESCRIPTION: SEC ID NO: 79:					
5	POTENDORS SECRETEDES ASSOCIAS SACISTITO AUGUSTOSI DAASSTOOD	6(
	CAGATUCOGA, (#38/976/1AGA CTCTFTCTCT (#ACAAT/93GA ACAGCOGACA GTGATGAGAT	120				
10	COCRECTE SEASOACE DA PROPETE DE CASASSA ACASSAS ACAS ACASSAS ACAS ACASSAS ACASSAS ACAS AC	181				
	GGCCAAGUTC CTGCTTCACCT GCTGCTTCTGC GCTYCGGCCCC CGGGCCACCC AGGCCAGGGG.	240				
	CAGCAUCUGG CTGCTGGTGG CCTCTTGGGT GATGCAGATC GTGCTTGGGGA TVTTGAGTG	300				
15	AGTOCIA MAA GEATTTITITI ACATOOMOSA OTACACOTO CICGROACCI COMMAGGORM	3€0				
	CATCINGATA GREGOTISTES CIGITECTOSC ORGASCITECT SCOTTICATIT ACGARAAC	42(
30	PACACATOR TECEACESTS SOCIATOR TOT CARRACTEST SCOREGINAL ACACESTED	480				
20	CATCGOD AND COMBABITIT GRAPHSARA TOTOGRAPAT GREETAGTOTT MUTACAACAC	54(
	TGCCTNCCCC ATCTCCAGCT CGAGTGAGTN GAACACTCCA GCCCCCACTC WAAGTCCAG	600				
25	AGAASTSAA AGASTACATS STOCKTO STOCKTO AGAASTSAA AGAASTSAA AGAASTSAA	660				
	AACCOTTING GOVATGETCT TEEGREETS SATTETESTS CTTCTGGEAT CTCTGGCCC	72 (
20	CAAASSAASA OOAQAQAAAA SQQAAACCAA COOTTQTAAQ ASQTOSTOAT GTCSGTGTTOT	78(
30	GTHGGAAGTG AGTGGAATCT AGCCATGCCT CTCCTGATTA MTAGTGCCTG GTGCTTCTG	84(
	ACCESSESTIC CONSCRICTS ACTIVITIES AASAACAS ACTISASEA ASAGETOT	9(:(
35	CARCAGONO ASTINICCTO COCCATANO NACIONAL CONTROLO SOCOTORIO DI CARCAGONI	960				
	ADTEMPERED COTTOCOLIER ACTICEPONTO DESCRICA TELECONACIAN INCIPATIAGO CO	1020				
40	TOTGALAATA AACKICAIG TIATIGTINN NAAAAAAAA AAAAAAAAA AATHTGGG	108				
40	GGCGCCCSTA CCCATIGSGC CTMMSGSSGN GGTTTAAAAT TAATGGGGGG GGTTTAAAAC	114				
	GGP:	114				
45						
	TO AN ALVERTAGE TO TE NO. CO					
50	(2) INFOFMATION FOR SEQ ID NO: 80					
50	(i) SEQUENCE CHARACTERISTICS (A) LENGTH: 557 base pairs					
	(B) TYPE: nucleic acic (C) STRANDEDNESS: double					
55	(D) TOPOLOGY: linear					
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 80:					
	GGCAGACAGC AGATGGCCTT GACACCAGCA GGGTGACATC CGCTATTGCT ACTTCTCTGC	6				
60	MCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	12				

WO 98/39446 PCT/US98/04482

	COCAGUTORO CATGATORAT COMEGUAÇA TOCUTOTURO TACOUTAGUTA CAGUTORO CA	180
5	CASCIMAGAM GAMPAGAM GAGAGATMAT CAMPAGCITAC CITITIACCOI GEMACTICAE	240
•	GETETRETTE CESATERES TECCTETETE RECOGNICET GGCAGGOTT GEGOTGCIV	300
	ADROCROACO SECACECER TOOTIETEED EEREEMECMA DISETSECTA DESCRIPA	360
10	TOOTTOOTAET CEEEEWAGEA DOETAGAACT ACATCTEAAA CEETAEWAEA ACCCEDOCCO	420
	GCAGCTTGAA CCTTTGACTT CTGACCCTCT CATCCTGGAT GTGTGTTGTT GGCACAGGAA	480
15	COCCOCCOCCO AACTITTATA ACATTTAAA ACAATTTAAA AAAAAAAA AAAAAAAA	540
1.	AAAAAAAA AANTCGA	551
20	(2) INFORMATION FOR SEQ ID NO: 81:	
	(i) SEQUENCE CHARACTERISTICS:	
25	(A) LENGTH: 795 base pairs (E) TYPE: nucleic acid	
	(C) STRANDEINESS: double (D) TOPOLOGY: linear	
	(xi) SEÇUENCE DESCRIPTION: SEÇ ID NO: 81:	
30	GCCGGGGGGAT TGTGGAGCGC GGGGGGGGC GGGGGCCT GGCCGGTGCT GTTGGGGCTG	€(
	CTBCTBBCBC TSTTAGTGCC GBBCBSTRBT GCCGCCAAGA CCGGTBCBGA CTCGTGACCT	120
35	GOGGGTOGGT GOTGAAGOTG OTCAATAOGO ACCACOGOT GOGGTGGACT OGCACGACAT	180
	CAAATAGGGA TOOGGCAGCG GCCAGCAATC GGTGACCGGC GTAGAGGCGT GGGACGACGACGC	240
	MAATAGOTAC TOGGGATCC GOGGCGGCTC GGAGGGCGGG TGCCCGCGCG GGTCCCCGGT	300
40	GOSCTSCESS CAGGESSTGA GSCTCACGEA TSISCTTACG GSCAAGAACY TSCACACGCA	360
	ASOSSOBA SESSTITUS TRADISPASS ASSASASS TETUSSOBAS ASSESSED DOSTITUS	47:0
45	GEGACIDAC CTGSACCTAT GEACAGTGES CTGTTGTA CAGCACTGE ACCTTGCGG	480
	TOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTO	540
-0	TACDATETAS AACOETEACO DETAGEORAD CTERAETADO ACESTECOT ACOESTECO	60
50	ACCTEGACET CTCCCGAGET GTGCTCCAC CACTTCTAC EGAGECTACC GGAGETGCC	660
	GATGAACTET GAGTGTGGG ATGAGTGGGT GGATGAGGTGG GGCGTCTGCA	720
55	GBBCCACTCT TGBCAGABAC TTTBBBTTTB TAGGBBTCCT CAAGTBCCTT TNTBATTAAA	780
	GAALSTIPST CTATG	79

12	;	INFORMATION	FOF.	SEÇ	IĽ	NC:	٤2	:
----	---	-------------	------	-----	----	-----	----	---

<u> </u>	(1) SEQUENCE CHARACTERISTICS: (A LENGTH: 1324 base pairs (B: TYPE: hitleic acid (C: STRANDEDNESS: double (D: TOPOLOGY: linear	
l C	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 82:	
	NAGGOTTIAA AGOGOTTACO CTXXOTGOAG GTGAGCAGTG GTGIGTGAGA GOCAGGOGT	6%
1 5	COTOTOCONG COCACTOAGI GGUAACACO GGGAGGTTT TTGICCITTG TGGAGGCTCA	120
15	GCAGTTCCCT CTTTCAGAAC TCACTGCCAA GAGCCCTGAA CAGGAGCCAC CATGCAGTG	18.
	TODAGOTTOA TIARGACCAI GANGAINCE CIDONATTIANO TOAGNITION GEORGEA	24.
20	GIPTISTEGG CACTOSTACAT CITARA ACTORA ATTOCATOR CATTOCTTOCT CAAGATOTTOT	30:
	OPECCES TO TOTAL TEST CATCALACTE TEST CATCALACTES CATC	36
25	GROSTIGUES TOTITISCUOT USTUTUCONE GEOUTSCUATES GUESCAAGAC USAGAGCAAG	421
4 -	TERROCCTOR TGACETICTI CTICATCCTE CTCCTCATCT TEATHECTGA GGTTGCAGCT	480
	GITGTGFTCG CCTTGFTGIA CACCACAATS GCTGAGCACT TCCTGACGTI GCTGFTAGTS	541
30	CODECCATICA AGAAAGATTA TGETTOCCAE GAAGACTTCA CTCAAGDERG GAACACNACC	€(
	PRODUCTO APPARATUTA DEDATATOAS DOADTTOEDE DIOETDAADT DEDEAARIA	664
35	CARCACACA DADTECARACA ETAGODOTO TOCCOCCOCATO ACARCACACACACOCC	72:
-/-	AATSAAACCT GCACCAAGCA AAABBTTCAC GACCAAAAAG TAGABBTTTG CTTCAATCAS	780
	CITITESTATG ACATOCOBAAC TWATHCAGTC ACCGIGGGIS CICTRECAGO INGGAATINGS	84.
4 0	GGCCTCGAGC TEGCCAT GATTETKTCC ATGTATCTGT ACTGCAATCT ACAATAAGTC	901
	CACTTCTGCC TOTGCCACTA CTSCTGCCAC ATGGGAACTS TGAASAGSCA CCCTGGCAAG	960
45	CAGCAGTGAT TOTAGAGAG GACAGGACAG ACAGAGCTTOTAGACAGTGAG AATGAGACAG	102
7.	CCCTTTCTGC TCCAGACTTS GGCCIAGATA GGGACCACTC CTTTTAGCGA TSCCTGACTI	1080
	TOOTTOCATT GOTHYGOTGGA TGHOTGHAHG GOATTOCAGA GOOTOTAAHG TAGOCAGTTO	1141
50	THETTIGOCCAT THOCOCCAGTO TATTARACCO TIGATATGOC CUCLAGROCT ASTRGIGATO	1260
	COAGTGOTOT ACTOGGGGAT GAGAGAAAGG CATTTTATAG COTGGGCATA AGAGAAATCA	1260
55	GUAGAGUCTO TUGUTGGATG TOTABAAGGO ACTICAAAAT GUATAAACCI GITACAATGI	132(
-/-	TAAJ.	1324

(1) INFORMATION FOR SEC ID NO: 83:

(i) SEQUENCE CHARACTERISTICS:

5		(E) TYF	GTH: 1494 b E: nucleic ANDEENESS: CLOGY: line	acid doubl€			
10	(xi) SEĮUENCE	DESCRIPTION	: SEQ ID NO	: 8 3:		
10	CTCAGGCTTC	TGTCTCACTT	TTOOGGGGGG	GEGATTAGGG	CAAGGAGGGC	ATGABGGACT	€€
	GICTCTCCCT	AAAACCCAGA	CONCRETACO	CCACTCAGTT	CTTCTTCATC	ADTOCTOCTO	12:
15	TCTTCATTGC	TGAGETTECA	GOTGOTGTGG	TOGCCTTGGT	GTACACCACA	ATGGTGAGAC	180
15 TOTTCATTGO TGAGGITTGCA GOTGOTGTGG TOGCCTTGGT GTACACCACA ATGGTGAGAC ACTGGGATGG AGGAAGGGAA GAAGATTGGG CAAAACCCTG GGAGTGGGCT GTGGCCTGTG AATGGCCACC TTCTGTACCA GOCCCTAAAC ACTGGCCTGC CTCACCCAGG CTGAGCACTT CCTGACGTTG CTGGTAGTGC CTGCCATCAA GAAAGATTAT GGTTCCCAGG AAGACTTCAC TCAAGTGTGG AACACCACCA TGAAAGGGGT AAGGTTGGCT GGGGGGAGGGT TTAGGGTGGA 25 GAGAAAGAAG CAAGGCCCCA CCTCCACCCT CATCTTGTCT CCAGGTCAAG TGCTGTGGCT TCACCAACTA TACGGATTTT GAGGACTACA CCTACTTCAA AGAGAACAGT GCCTTTCCCC CATTGTGTTG CAATGACAAC GTCACCCAAC ACAGCCCAAT GAAACCTGGA CCAAGGAAAA GGCTCACGAC CNAAAARTAN AGGTGTGGC TGGCATGAGT GGGTGGGAGC TGTTTTCATG GCCTCAGGAGT GGCAAACGGG GATGGGGTGC CAACTATAAA TGCTGTTTTC 35 TCTTCCYGAA GGGTTGCTTC AATCAGCTTT TGTATGACAT CCGAACTAAT GCAGTCACCG	240						
20	AATG3CCACC	TTCTCTACCA	GCCCCTAAAC	ACTGGCCTGC	CTCACCCAGG	CTGAGCACTT	300
20	CCTGACGTTG	CTGGTAGTGC	CTGCCATCAA	GAAAGATTAT	GSTTCCCAGS	AAGACTTCAC	360
	TCAAGTGTGG	ASSASSSAA	TGAAAGGGGT	AAGGTTGGCT	GBBBBBBBTT	TTAGGGTGGA	420
25	GAGAAAGAAG	CAAGGCCCCA	COTOCACCOT	CATCTTGTCT	CCAGCTCAAG	TECTETECT	4 80
	TCACCAACTA	TACGGATTTT	GA/FBACTCAC	CCTACTTCAA	AGAGAACAGT	GOOTTTCCCC	547
30	CATTOTOTO	CAATGACAAC	CASCOCACTO	ACAGCCGAAT	GAAACCTYGCA	CCAAGCAAAA	€(+)
30	GGCTCACSAC	CNAAAARTAN	AGSTSTGGGC	TGGCATGAGT	GGGTGGGGAC	TGTTTTCATG	660
	GCCTCAGAGT	GGCAAACGG3	GATGGGAGTA	GGGCAGCTGC	CAACTATAAA	TGCTCTTTIV	72%
35	TCTTCCYGAA	GGGTTGCTTC	AATCAGCTTT	TGTATGACAT	CCGAACTAAT	GCAGTCACCG	780
	TGGGTGT	GGCAGCTYGGA	ATTAGGGGGC	TCGAGGTAAG	CAGATSAGGA	GCTGGGACTG	841
40	GGACATGGGC	ATGAGACCAG	GGCTGCTCAA	CCCATCTGAG	GCCTCTCTG3	AGGAAACAGA	900
	CTTCTAACTG	GGCCTCAGGT	AGRETETETS	TYGGGACAGGC	TTCAGGATCC	CTATCATGTI	960
	CCCTCATCTC	TCCCTGTTCC	ADDITOTOOT	GETGGCTGCC	ATGATTGTGT	CCATGTATCI	1020
45	GTACTGCAAT	CTACAATAAG	TOCACTTOTG	CCTCTGCCAC	TACTGCTGCC	ACATGGGAAC	1080
	TGTGAAGAGG	CACCCTGGCA	AGCAGCAGTG	ATTGGGGGAG	GGGACAGGAT	CTAACAATGT	1140
50	CACTTG33CC	AGAATGBACC	TBOCCTTTCT	GITCCAGACT	TGGGGGTAGA	TAGGGACCAC	1200
	TCCTTTTAGC	GATGCCTGAC	TTFOCTTCCA	TTGETGESTG	GATGGGTAGG	GGGCATTCCA	1260
	GAGCCTCTAA	GGTAGCCAGT	TOTGTTGCCC	ATTCCCCCAG	TOTATIAAAC	CCTTGATATG	1320
55	CCCCCTAGGC	CTAGTGGTGA	TOTOLITIET	CTACT/GGGGG	ATGAGAGAAA	GECATTITAT	1380
	AGCCTGGGCA	TAAGTGAAAT	CAGCAGAGCC	TCTGGGTGGA	TGTGTAGAAG	GCACTTCAAA	1440

ATGCATAAAC CTGTTACAAT GTTAAAAAAA AAAAAAAAA AACTCGACTC TGCC 1494

5	(1) INFORMA	KTION FOR SE	5Ç ID NO: 84	G :			
10	(±)	(E) TYP: (C) STR	HARACTERIST GTH: 1285 b E: nucleic ANDEDNESS: CLOGY: line	ase pair: ació doubl:			
	(xi)	SEQUENCE I	DESCRIPTION	: SEÇ ID NC	: 84.		
15	GCTACCTYGGT	TACOCACCOT	GGGAACGAGG	CCCTGGGGG	GGAGTTGCTT	CTGCTCCTGA.	ϵ
15	TGCAGTTOOT	DASTACCETO	TTOCTECGAG	SGAAOCCAOG	6900 00A 00099	CISCICICIO	12
	AGATGCGCAT	TOACOTSITE	CCCTCCATGA	ACCCTRGATIGS	CTATGAGATC	GCCTACCAC:	18
20	GGGGTTCAGA	FORSTESECT	TGGGGCGARG	GDCGCTGGAA	CAACCAGAGC	ATOGATOTTA	24
	PTTAATADDA	THEOMSAMOTO	DACCACCAG	TEPAAGE	ACAGGACGAT	GGGAAGSTG	3(
25	CCCACATOGT	CCCCAACCAT	CACCTGCCAT	TGCCCACTTA	CTACACCCTG	CCCAATGCCA	36
23	CCGTG3CTCC	TGAAACGTG3	GCAGTAATCA	AGTGGATGAA	GOGGATOCCC	TTTGTGCTAA	42
	CTGCCAACCT	CCACGGGGGG	GAGCTCGTGG	TSTCCTACCC	ATTOGAÇATG	ACTCGCACC:	48
30	CGT/333CTG?	CCGTGAGITT	ACGCCCACAC	CAGATGATGC	TETETTTEGE	DESCTOAGOA	54
	CTGTCTATGC	TAGCAGTAAT	CTGGCCATGC	AGGACACCAG	COGCCGACCC	TGCCACAGCC	€0
25	AGGACTTCTC	CSPROADSGO	AACATCATCA	ACGGGGGCYTG	ACINGGCACA	OBSTOCOCO	66
35	GANGCATGAA	COAC TTYAECT	TACCTACACA	CCAACTGCTT	TGAGGTCACT	GTGGAGCTG:	72
	AADAETETDE	OAC CCCCTTD	GAGAATGAAT	TGCCCCAGGA	GTGGGAGAAC	AACAAAGAGA	7⊦
40	CACTOOTOOO	CTACCTGGAG	CAGGTGCGCA	TGGGGCATTGC	AGGAGTGGTG	AG93ACAAG3	84
	ACACGGAGCT	ТЭЭЗАТТЭСТ	GACGCTGTCA	ASSETSOCERT	TYGGGATTAAC	CATGACGTGA	90
	CCACGGCGTG	GGGGGGAT	TATTGGGGTG	TGCTGACCCC	AGGGGACTAC	ATGGTGACT	94
45	CCAGTKCCGA	GGGCTACCAT	TCAGTGACAC	GGAACTGTCG	GGTCACCTTT	GAAGAGGGCC	102
	001110000013	CAATTT 19T3	CE CAT CAAGA	CTOCCAAACA	JAPPOTFOGO	GAGCTGCTGC	108
50	CAGCTGGGGC	CAAGET GCCC	CCPGACCTTC	GCAGGCGCTT	ATCEECEAEE	AGGGGACAG/	114
	AGGATTGATA	OCNSOGGTTI	AAGAGTCCTA	GGGCARGCTG	GACCTSTCAA	GACGGGAAGS	120
	GGAAGAGTAG	AGAGGGAGGG	ACAAAGTGAG	GAAAAGGTGC	TCATTAAAGC	TACCGGGCAC	12 6
55	СТТААААААА	AAAAAAAA	AAAAA				128

	(2) INFORMATION FOR SEC LD NO: 85:	
5	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 394 base pairs (E) TYPE: nucleic acic. (C) SIRANDEDNESS: double (D) TOPOLOGY: linear	
0	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:	
0	GCGCGCTCTA GGAACTAGTG GATYCCCCCGG GNCTYCCAGGT GT/GGAGTGGG CCATCGTAAA	€(
	TAGTATCTGT GCATAAGSTS GTTGTGCGAT AAATGAGTTA ATGTATGCAA AGCCCTTGGC	124
15	CCAGAGCOGG CGCAGAGCAT TGTGTAAGTS CTGGCAGGGG TCATGATGGA GATATCATGI	180
	CTCCTCTTRI TGATTCAGGA TICTGATGAG ATGGAGSATS GC-CCTGGGGT TCAGGATTAG	24
30	GCCTTGAGGC ACTGCTCCAG CCTCCTTTGT GGGCCCTGTC ACCCTTGGCT TCATCGGGCC	30.
20	GTARCAAGTC TCCCCTCICC CACTYTGCAG CAGARGTGTT CAAGAACTGC CTGCTCACGG	36.
	TICGIGITCI GCAAGGCCAT CGCCTAACCT CTAA	394
25		
	TO NO. 96	
20	(2) INFORMATION FOR SEC ID NC: 86:	
30 35	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1925 base pairs (E) TYPE: nucleic acid (C) STRANDEDNESS: double (E) TOPOLOGY: linear	
<i>.</i> ,	(xi) SEQUENCE DESCRIPTION: SEQ 1D NO: 86:	
	AGTGAAGGGA GCTGGTCGTG CGACTGGGCT TCGGGCCCTG TGCCAGAGGA GCANGCCTTC	€(
40	CTGAGCAGGA GGAAGCAGGT (GTTGGCCGCG GCCTTGAGGC A: PGCCCTGCA GCTGGATGGA	12(
	GACCTGCAGG AGGATGAGAT CCCAGTGGTA GCTATTATGG CCACTGGTGG TGGGATCCGG	18(
45	GCAATGACTT CCCTGTATGG GCAGCTGGCT GGCCTGAAGG AGCTGGGCCT CTTGGATTG	24(
4.	KTCTCCTACA TCACCGGGGCCC CTCGGGCCTCC ACCTGGGCCT TFGCCAACCT TTATAAGGAC	300
	CCAGAGTGGT CTCAGAAGAA CTTGGCAGGG CCCACTGAGT TGCTGAAGAC CCAGGTGACC	36(
50	AAGAACAAGC TGGGTGTGCT GGCCCCCAGC CAGCTGCAGC GGTACCGGCA GGAGCTGGCC	42(
	GAGCGTECCC GCTTGEGCTA CICAAGCTGC TTCACCAACC TETGGGCCCT CATCAACGAG	48 (
55	GCGCTGCTGC ATGATGAGCC CCATGATCAC AAGCTCTCAG ATCAACGGGA GGCCCTGAGI	540
55	CATGGCCAGA ACCCTCHGCC CATCTACTGT GCCCTCAACA (CAAARGGCA GAGCCTGACC	600
	ACTITIGAAT TIGGGGAGIG GIGGGAGIIC TOTICCIACG AGGIGGGIIT COCCAAGIAC	éé(
60	GOGGOCTTOA TOCCOTOTGA GOTOTTTGGO TOCGAGTTOT TYATGGGGCA GOTGATGAAG	720

	FORGINGING FORGOOOGEN CIRCUICINE GFFEENWEIGT (APPOEMFULL GNAINAGFFAG.	/ 1-1
5	AACCTCUAGS ACASTTATA CTOS KOTICA GAGCCOAGAC ACTICIOAGA COGSTGOST	84
•	ACCARACTERS (CARCOTOGA CARGARCIAS STOCCOSTIC TORAGATAGA AGAACCASC	9()
	TORANGARYON GCAGAATARC NYGAGTITTIN ACCGATOTTO TYRACOTGROG TYCACTGRO	٠ ١ ١
10	CARSCLACAC ASAACTES STORERSSON CATTOCIADA AAGACTACTE TCAGGACIACE	1010
	CASTISSTANA (ATSEAA FO TACCACTISE GATGREETTOC SCAASCASCI GAGACCSSSC	1(8)
15	TODOCETICO OTTERACIONAL ACTROCTORIO OSETTETAREN TORICORES OTTERACIONAL ACTROCAGENO	114
13	DEMARGNACO IDDAAGA CHA KEMDAGEDO IDAGED DEGI AKEENDO AGGACENDO AGGACO	110
	TOGASSCRIPTO DOSTAFFASA (REALBASCRIPTOTTROCORRER) INSTRUMENTO DOSTAFFASASCRIPTOTTROCORRERO DOSTAFFASASC	126
20	ATOTHOR SECOND TECHNOLOGIC TE AEFRICIO RESCRIPCIO AE AAROOND AR AAROOND ARCONDECTE OF THE ATOTHOR ARCONDECTE OTHER ARCONDECTE OF THE ATOTHOR ARCONDECTE OF THE ATOTHOR ARCONDE	132
	-AGERBOOTTO OTOAGER ACT FETCHONING DAIGHTERS DETOCCORAS FORCESTOA	138
25	TACTOSPINO CONSESTORS GOSSACACOC GASSAGSOS CAGOTOSPISA GETSAACOTA	144
25	TOTTOAT 193 ACTOTOCOTA COACTACAGS AARRIGACOT ACARCOAGSA GSACGT SSAC	150
	AAGCTUCTUC ACUTSACACA TOACAATETO TURCAACAADO AUGAGCAGOT GOTGGAGGOT	156
30	CTECOSCIENCE CASTSCASIG GASSCESCAS CECASGOCCO ACTGATISCO GESSCOCCIV	162
	CACCTOCTET DAAEDETSSA CETTSACITS IDEETCOCTT ACTTACHTTO AATOCODADO	168
35	CASTIGETICA GASCITORIG CTICAGRIEGO ACTRICOCAR GRICOAGGIT GAGIGGETIGA	174
55	AGCTOCCTIS COUCTOAGOA STITSCAGIS GREVAAGGO FACCAGOOGA TITGTOTAL	180
	ASCITRATE ASTIDATES DETECTIONE TETRICORTS TOUSPOONS DAAAADOOAD	186
40	CACTTGATAC ATGACAGACT CATACAAATG TYAGGGGCTG AGAAAAAAA AAAAAAAAA	192
	CTCGA	192
45		
73	(2) INFORMATION FOR SEQ ID NO: 87:	
	(i) SEQUENCE CHAFACTERISTICS:	
50	(A) LENGTH: 1818 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
55	(xi) SEQUENCE DESCRIPTION: SEC 1D NO: 87:	
	CCGGGCCCC CCNCGNGKTT TTTTTTTTTT TTTTTTTTTX TATGAGTCTG THATGIATCA	6
60	AGIX-CTCCAA CLACTCAAGG TAGCGCAGAA GUGAAAACAG GCACAGGCCG GGGGGTTTTV	12
00		

	GGTGATTACA	CAAATYBBGCT	TGGTCTCCTI	ACCCCACTGC	AAACTGCTGA	GGCGCAAGGS	187
	AGCTCCCAGC	CCTCAGCCTG	GACCCI HOUA	CAGINGCCAIC	TGAGCCCGAG	GCTCTGNAAU	24.
5	CACTGCCTGA	TGACAGTTCC	CACCTGLAAC	TCAGCAGCCA	GGGAATGAAT	GAGASTTAGU	300
	GGTGGCAGGG	ADDEPORTOR	TOAGTGRAGG	CTGCGCTGCC	GICTICGCTG	CACTGCCTGS	360
10	CGCAGAGECT	COMBONAGERO	CTC TIVESTIF	TIGGAGACAT	TYPIAATGTGT	CAGGYACAG	420
10	AGITTGTHICA	CCTCCTCCTG	GCTGTAGGTC	ACCTICGTGI	AGTGGTAGGG	AGAGTVCGAT	480
	GAAGACAGGT	COCOTOGASI	AGCTGCCGC	TOTTOGGGTG	TOOGOOGGAO	CCCAGREGCC	540
15	GAGTACTCCC	GBAAGBAGTC	GCTGAUCAGA	GBAAAGTGCA	GCACCGCAG3	GGCTCCGGGG	€00
	CAGGTG793T	CGGAGAAGGT	GTGGCACTCC	CGAGGCTGGA	GITGOTOTTO	GBBGITGBB	€60
20	GAGATGUFTS	GGAACGGGAT	CCCCTGCTCC	TGGCAGAACC	GGCCCAGGAG	CTGCAACTG	720
20	TGGAAGGCTC	CSTGSAGSTT	GTAGTCCAAT	GACAGGATGA	GETECACGTE	COGAGTGGG1	780
	TGCAGGAGGG	GCAGGCAGCT	GGTATTGATG	AGGTAGCCAA	CATCCAGCAG	GCACAGGTG3	840
25	GGTTCCGAGG	GTGTCAGCTG	GTTGGGGAGC	COATCCAGAG	TGGTAGCTTT	CCATGTGGAG	900
	aastgagsat	GITGAAAGTA	GTCTT131/G3	AAATGGAGGC	CACGCAGGAA	ATTATGTGTG	560
30	ADINESTRIS	GIGGACGCCA	CGTCAGAAGA	TEGGTGAAAA	ACTCAGCTAT	neneceseer	1020
50	GTTGAGGGTG	GTTCTTCTAT	CTTCAGAAGE	GGGACCTGCT	CCTTGTCCAG	GTTGGCTGG	1080
	TTCCTGACCC	AGCGGTCCCA	GAACT/GGCT/G	GECTETGAGG	CCCAGTATAA	GOTGTOCTGG	114(
35	agettgettg	CATACAGGTT	GCTCCAGATA	COTTOTAAGA	AGCAGATGCG	GBACTCAGGA	1200
	AGTCTTTTCA	TCAGCTGCCC	CATAAA GAAC	TOGGAGCCAA	AGAGCTCAGA	GBGBATGAAB	1260
40	GCCCCGTACT	TGFFGAAGCC	GACCTCGTAG	GGAGAGAACT	CGCACCACTC	COCAAATTCA	1320
.0	AAAGTGGTCA	GESTOTGOOD	TTTGGTGTTG	AGGGCACAGT	AGATGGGCAG	AGGSTTCTGS	1380
	CCATGACTCA	GBBCCTCCCG	TTGATCTGAG	AGCTTGTGAT	CATGGGGCTC	ATCATGCAGT	1440
45	AGCGCCTCGT	TGATGAGGGC	CCACAGGTTG	GTGAAGCAGC	TTGGGTAGCC	CAAGCGGGCA.	1500
	ADDERDTEND	GCICCTGCCG	GTACCGCTGC	**************************************	GGGCCAGCAC	ACCCAGCTTG	1560
50	TTCTTCTTCA	CCTGGGTCTT	CAGCAACTCA	. GIGGGCCCTG	CCAGGTCCTI	CTGAGACCAT	1620
30	TCTGGGTCCT	YATAAAGGTT	GGCCAAGGCC	CAGGTGGAGC	CCGAGGCCC	GGTGATGTAG	1680
	GAGAC(+)AAT	· CCAAGAGGCC	CCAGCICCTI	TCAGGCCAGC	CAGCTGCCCA	. TACAGGGAAG	1740
55	TCATIGCCCG	GATCCCACCA	CCAGI IGCO	. TAATAGCTAC	: CACTGGGATC	TCATTCTCCT	1800
	GCAGGTCTCC	ATCCAGCI					1816

Light Contraction

	(L) INFORMATION FOR SECTIONS: 86:	
5	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 539 base pairs (E) TYPE: nucleic acid (C: STRANTEDNESS: double (D) TOPOLOGY: linear	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 88:	
	AGGGTAATTA ATATGAAGTG CAAAAAGTTG AATGTTCCAG TCTAAAAGGC AGTGGGAGAA	€ (
15	ATTACATAGO ATGAAATAA TAAAANGAAY TOTTATTAAT GAGAACGAGG YTOTTGCAGI	120
1.	GECAASTYCT GCTGETCACC CGATGGEGAT GEBAGCCTTT CAAGCTYTTT TYTEGGETAAT	181
	ACCICACACT TOCCACCOTON GENERAL PETECACACT TOCCACACTOR TOCCACACTOR	24(
20	CATCITOAAAG CITCATGSTG ACGCAGAGAGT CITSTTGAAGA TOACAGGSTC CITCGCTTGCA	300
	COADRAG DEAASTUCET COTOACOCOA TESTUCACITA COATECUTO DOATACERT	360
25	ADAACTOOOC TOTOAGTTOT COTTONATIAA ITAERATIAAA ARTORARAA TAERRATTA	420
	POTOTACOAA AGGAAACOGO TTTTOAACAD TOGTGGCOAA AASTGTEENT TGOTTGOTGT	480
	ACCETOARD OFFCCEAETO TOCCOPTION CEAUTAINA AEROAETE TO AAETCRUUUD	534
30		
	(2) INFORMATION FOR SEQ ID NO: 89.	
35 40	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 855 hase pairs (E) TYPH: nucleic acid (C) STRANGEDNESS: double (D) TOPCLOGY: linear 	
40	(xi) SEQUENCE LESCRIPTION: SEQ ID NO: 89:	
	CCTCTGCCCA GGCCGCACCC GAGCTCAGGC TCGTGCCCAC CCACCAAGTI CCAGTGCCGC	61
45	ACCRECATE ASERTCANCE ACADOMICEO SERCOACIOO COERECCENT TOEOTRACOA	120
	PROAGOGATG AGRAGGASTS CARRACTERS CONTETACOC AGAAARREDA ATGCCCACOG	180
50	CAADAGHCAA EERRAUCTSD TOADTDACTE GEDOOACHTS COCEUTOCOOU DOSEUTOCOOO	240
	CAGETOGOA ACTEDOTES COLEGACEATO CENTOCOE TO BOCEACOTOA ACCOMENTA	300
	GATGACTGCA TTCCACTCAC GTGCCGCCGGCGACCCCCCGAGACTCCCCCGACTCCAGC	360
55	SACSAGCTOS GCTGTSBAAC CAATSAGATE CTCCCGGAAG GRGATGCCAC AACCATGGGG	420
	CCCCCTGTGA CCCTGGAGAG TGICACCTCT CTCAGGAATG CCACAACCAT GGGGCCCC	480
	GTGAACCOTG GMGAGTGTOC COTOTGTOGG GAATGCCACA TOOTGCTOTG COGGAGACCA	540

	COORACOTO ACTORTORO TODICORACON TATTORORNIA COORTOAACO CEARDOTOTO	€00
	CARROCEUR SCESEARRAD DORARGOTOR RECOTETET COTESTANA SCESCAVER	660
5	CONTROLTO ASASASASAA AOTOTONTO ICCCUSASSA ASTACCOSTO SUCAUTSESD	720
	ADAEGAACHA ASEMTATROD EDTOCCEAT TOACTROCACO ACCETTOACO	780
10	CORRECTION CEASITORARA FROTTCOORA (CORLOCUES ORTERSITADO DECLARIONADO	84(
10	ACDSSAACTT CGAAC	851
15		
	(2) INFORMATION FOR SEC ID NO: 90:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 628 base pairs	
20	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 90:	
	ADDRDAERTA ERETERRIDO DIDDIDARDO DRABILITED ETIDDIDDIDI DETDIDARRA	60
	ASOTOTOTO ASERTETTA OTOSOTAAGACACACA AGACTOSO ASERTESTOS	120
30	CAGATTCACA AGGACTTAAT CT393TF3C GCGGGGCCT GTCAGAT3AG CATGTTGAG	180
	TGATATCTST TCTAGCCCAS CAAGCAGTA AGCTAACCTC TGACCCCACT GATATTCCTG	240
35	TGSTGTCT AGAATCAGAT AATGSGAACA TTATGATCCA GAAACACGAT GGCATCACGS	300
	TGSCAGTGCA CAAAATGGCC TCTTSATGCT CATATCTGTT CTTCAGCAGC CTGTCATAGG	360
	AACTGATCC TACCTATGTT AATTACCTTA TAGAACTACT AAAGTTCCAG TAGTTAGGCC	420
40	ATTCATTTAA TGTGCATTAG GEACTTTTCT GTTTATTTAA GAGTCAATTG CTTTCTAATG	480
	CTCLATGEAC CGACTATCAA GATATTASTA AGAAAGGATC ATGTTTTGAA GCAGCAGGTC	540
45	CAGSTCACTT TGTATATAGA ATTTTGCTGT ATTCAATAAA TCTGTTTGGA GGNAAAAAAA	600
7.0	AAAAAAAAA AAMTSGAGGG COGAAGCI	628
5 0		
	(2) INFORMATION FOR SEQ ID NO: 91:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1053 base pairs	
55	(B) TYPE: nucleic acid	
	(C) STRANLEDNESS: double (D) TCPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:	
60		

	COORDICED REPORTABLY ACADEMENTA RABBARARA ANABOMERAS BEDITATIONS	€ /
	PARACACIEM INCUCANDA ACTORNOCO CONCORARY INCRESIONA ACROSTORED	1.
5	CTGTOCGSAG (CCCACAACAC CACAGTGTTC LAGGGGCGTGG (GGGCCCAGTC CCTGCAGGT.	181
	PTOPACCECC ETERTICISES ACRIVERACION SOTINACIONAME INSCITUARITA IDDOCUNTOT	24(
10	PTTOTTOTOTOS TOFOTYTOA AOAGOADEA CTFFFTONS ACCEPTADIO DOAAGADAD	30
10	COLOROTORA ESTRESTON ATAGORADA OTACOGRADA DASSENTARASE TODACTORADO	360
	TACCTCCQAQ ACCQTGACCA TOTCTQAQCQ TAQCACCCA ACATCTAAAGQ CQCCCCCATTA	42(
15	GGCASTNAGS CTSACACTYT CARSAAPSTC CTRFTGSAGG TGCTGGCAGA CCCCCTGGAT	48
	CACCERSANTS CTESSAGATES STREETINECCE SCREAGETONG AGAGETTEGA GEATGCCCAT	54:
20	GTGGAGCACA, GCATCTCCAG GARCCTCTTS GAAGGAGAAA TCCCCTTTCCC ACCCACTTCC	600
20	REMOTEDINGO DASCRAGAS DITABANCHA DEDITETETAD STECKERTOTE SOCIATEDOTA	61.
	GCTGCAGCCT GGCATVGACA GAAGCCAGGG ACACATCCAC CCAGTGAACT GGACTGTGGC	724
25	CATGACCCAG GGTATCAGCT CCAAACTCTG CCAGGGCTGA GAGACACGTG AAGGAAGATG	78(
	ACCOUNTACTA DOTTORADO DACORDORO DACORDAR ASCACOLARA ANDREAS ANDREAS ANDREAS ANDREAS AND AND ANDREAS AND ANDREAS AND ANDREAS AND ANDREAS AND ANDREAS AND AND ANDREAS AND ANDREAS AND ANDREAS AND ANDREAS AND ANDREAS AND AND ANDREAS AND ANDREAS AND ANDREAS AND ANDREAS AND ANDREAS AND AND ANDREAS AND ANDREAS AND ANDREAS AND ANDREAS AND ANDREAS AND AND ANDREAS AND AND ANDREAS AND AND ANDREAS AND ANDREAS AND ANDREAS AND ANDREAS AN	840
30	CTTGGCCACC AGGACTCCTT GTTCTGCTCT GGCAAGAGAC TACTCTGCCT GAACACTGCT	90.
30	TCTCCTGGAC CCTGGAAGGA GGGACTGGTT GAGGGAGTGG GGAGGTGGTA AGAACACCTX	961
	ACAACTTCTG AATATTGSAC ATTTTAAACA CTTACAAATA AATCCAAGAC TGTCATATTT	102(
35	AAAAAAAA AAAAAAAAA AACNOGAGGG GGC	1055
40	(2) INFORMATION FOR SEQ ID NO: 92:	
,,		
	(i) SEMMENCE CHARACTERISTICS: (A) LENGTH, 1075 base pair:	
	(E) TYPE: nucleic acid	
45	(C) FTRANDHDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION SEQ ID NO: 91	
50	correction carried on proceduration moracecarra careacturer macariations	€.
JU	GCACGAGCCT GATCCICTCT TTTCTGCAGT TCAAGGGAAA GACGAGATCT TGCACAAGGC	
	ACTICINGTIO TGCCCTTGGC TGCGGAGAGGG TGGCATGGAG CCTCTCCGGC TGCTCATCTT	12:
55	ACTOTTTSTO ACAGAGOTGT COGGAGGOGA CAACACOACA GTGTICCAGG GGGTGGG	18
~~	CCAGTCCCTS CARSTSTCTT GCCCCTATSA CTCCATSAAG CACTGGGGSA GSCGCAAGGC	240
	CTGGTGCCGC CAGCTGGGAG AGAAGGGCCC ATGCCAGCGT GTGGTCAGCA CGCACAACTT	30(

60 GTGGCTGCTG TCCTTCCTGA GGAGGTGGAA TGGGAGCACA GCCATCACAG ACGATACCCT

60

Section 3.

	GBGTGGEACT	ATTACONAUTO	OGOTGOGGAA	CODAADATOT	CATGAT9033	GTOTOTACCA	42
5	GTGCCAGAGC	ACCOUNTACIONO	GTSAGGCTGA	CACCCTCAGG	AAGGTCCTGG	TGGAGGTGC	48
-	GGCAGA:0000	CTGGATCACC	GBGATGCTGG	AGATOTOTGG	MICCOCGGGG	AGTCTGAGAG	54
	CTTCGAGGAT	eccomaras	AGCACAGCAT	CTCCAGGAGC	CDCIDGGAAG	GABAAATOO	€¢
10	CTICCCACCC	ACCROCCACCO	TICTCCTCCT	GROCIGIATO	TTTCTCATCA	AGATTCTAG"	€6
	AGCCAGCGCC	GTGT19999513	CAGCCTGGGA	ФЭЗАСАЗААЗ	CCAGGGACAC	ATCCACCCAC-	72
15	TGAACTGGAC	MESSECCATE	ACCCAGGGTA	TCAGCTCCAA	ACTCTGCCAG	GGCTGAGAGA	78
••	CACGTGAA33	AAUMGATGE	CEAAAAEEA	CCAGGAGAAG	POCACCACC	GACCAGCCCA.	84
	GCCT3CATAC	STTOAGGGGT	GCCACCAGGA	CTCCTTGTTC	ACDOTOTOET	AGAGACTACC	90
20	CTGCCTGAAC	Y 31.33444343	CTGGACCCTG	GAAGCAGGAA	CTGGTTGAGG	CASSISTEDAD	Ç.
	GTGGTAAGAA	CACCOGACAA	CTTCTGAATA	TTGGACATT	TAAACACTTA	CAAATAAAT	102
25	CAAGACTGTG	AAAATTTACA	AAAAAAAA	AAAAAAAAACN	OGAGGGGGEN	CCCGG	107
	(2) INFORMA	TION FOR SE	EQ ID NO: 93	3 :			
30	(i)	SEQUENCE CI	HARACTERIST:	ICS:			
			GTH: 2492 b E: nucleic	-			
35		·	ANDEDNESS: OLOGY: line				
	(xi)	SEQUENCE I	DESCRIPTION	: SEQ ID NO	: 93:		
40	TCCCGACTCA	GCTTCCCACC	CTGGGCTTTC	CEAGGTGCTK	TOGCOGCTGT	CCCCACCAC	6
40	GCAGCCATGA	TCTCCTTAAC	GBACACGCAG	AAAATTGGAA	TESEATTAAC	AGGATTTGGA	12
	GIGITTTTTCC	TETTETTTE	AATGATTCTC	TTTTTTGACA	AAGCACTACT	SECTATTOED	18
45	AATGTTTTAT	TTYSTAGCCGS	CTTGGCTTTT	GTAATINGGTT	TAGAAGAAC	ATTCAGATT:	24
	TTCTTCCAAA	TAAAATACAA	GAAAGCTACA	GGTTTTTTTC	TGGGTGT	ATTTGTAGT	30
50	CTIATTGGTT	GGCCTTTGAT	AGGCATGATC	TTCGAAATTT	ATGGATTTT	TOTOTTGTT:	36
50	AGGGGCTTCT	TWOOTGTOGT	TGTTGGCTTT	ATTAGAAGAG	TECCAGTOCT	TGGATCCCTY	42
	CTAAATTTAC	CTGGAATTAG	ATCATT PGTA	GATAAAGTTG	GABAAAGCAA	CAATATGGTA	4,8
55	TAACAACAAG	TIGAACTI FGAA	GACTCATTTA	AAATATNSTG	AATATTTA PT	AGTCATTTGA	54
	AGAATATTCA	COACAAAACT	AAATTACATG	AAASAGOTTIG	TAATGTTCTT	TACAGGAGTT	60
				CALAMMARICA			66

	TTCIACICAA	OAATOAAUTO	AAGAAGTCAG	CAAGGAAACT	GAGAGAGGTG	AAATCCATGC	72:
	TAADGATGCT	CZAGAAACCC	TTGAAGGCTA	ninenemanat	TYTTTOCAGAA	TGTGCGAAA*	783
5	COMMODERADIC	TTAGAGAACT	erestenere.	TOTOTITIEST	MANI YAMMA	AAGYTTCAG	84.
	AGCATTCCATA	CONTINCO	TTTTAGAAAI	GI CUACTGCA	Aaaaaceeta	TATTTCCACT	90
10	receptions	CTCTGGAAGT	GATGCATGAI.	TO CAST TOGA	TTSTSTCATT	TIAAAGTAT	96:
10	AAAACCAAGG	AAACCCCAAT	TATECATETAT	GGATTACTTI	AAATƏITITT	CATGGTTAAV.	1020
	ATAAAACTTC	TGTGTTTCTT	CTGAATCTTA	ATAITTTCAAA	GCCAGGTGAA	AATICTGAACT	1080
15	AGATATICTI	TGTTYGGAMTA	TGCAAAGGTC	POATTTTTACT	AACTTTTAGT	TACTAAATTA	1140
	TAGCTAAGTT	TTGT CAGCAG	CATACTCCFG	AAAGTCTCAI	ACTT-CTT/3GG	AGT CTG-10CL	120
20	OCTABOLATO	PARTATORE	ATTICATTIACS	TGTAAGTATT	TAACAAAAA	GCATTCTTG7	1260
20	COATGAATGA	AGIAGITTGI	TTCATAGCT	GUCUCATTSA	ATAGTATTAT	TGAAGATACT	1317
	AAATGATGCA	AACCAAATGG	ATTTTTTCCA	TGICATGATG	TAATTTTTCT	TriCTriCitie:	1380
25	TTTTTTTAA	ATTTTAGCAG	TGGCTTATTA	TTT/?TTTTTC	AAATTAAATA	ATAACTTTT:	1440
	ATTATETTA	CTTTAAGACA	TGTAACATGT	IAAAAGSTTA	AACTTATTOC	TGTTTTAAL	1500
30	GGGCI ATTICA	TTTAATCIGA	GTTTTCCITT	ATTITICAGCT	TTTTCCTAGC	PATAATAG	1560
	CATIAAGIAI	GACATATOCT	TCATATGATC	ACTEATETTG	AGTTAATTAG	CODATAGAA	1620
	AGTTCACGTG	CTAAAGTCAT	TTCACTGTAA	TAAACTGACT	RIGGTTTOTT	AAGAACATGA.	1680
35	CACTAAAAA	AAAGTGGTTT	TTTTCCACCO	TIGTIGATIA	TTAGACAGTA	GGAAATAGC'.	1740
	GTTTTCTTTA	GTTTTACAAG	ATGIGACA 30	TTTAGTGGTA	GATGTAGGGA	AACATTTCAA:	1800
40	CAGCCATAGT	ACTATTISTT	TTACCA::T3A	TYPGDACTYTY	TIGITITITI	AACAGT DGCA.	1860
	AAGCTTTTTA	ATGCATAAAA	AS TTAATATO	. AAT TTGTNGGT	ATTTATTTAC	AAACATGTC:	1920
	ACAAAAATAG	ATTACAGITT	TITTAITITA	TAGTTAAATC	TCTTAATACA	CAGA/ENAAC'.	1980
45	CCCAATCTTG	AAATOTAOTO	TAAGGAAAGA	CTT GGTGTAT	AGTGTGATGG	TTTAGTCTTA	2040
	AGGATI AAGA	. CATTTTTGGT	ACTTGCATTI	GACTTACGAT	GTATCTGT 3A	. ASATEGGATG	2100
50	ATATTGACAA	. ATGGAGACTC	CTACCTGAAT	AGTTAANGGA	ATAATAAGAG	GTIACIGTY	2160
• -	TGICTAAIGT	TCTTCAAAAA	. AGTAATATCC	TCACTT 3GAG	AGTGTCAAAT	ACATACTTK-	2220
	AGGATTGACT	ODAATATATT	FATE FOODER	AAMTOTGTTA	CACATATTT	TGACCCATAT	2280
55	TATTTACAAT	GICTTGATAA	TTCTACETTS	TTAGAGCAAG	AATAGTATCI	GOTAATGTAA.	2340
	GGGACATCTC	TATTTAACTC	CTTTGTAGAS	ATGAATTTCT	ATCAAAAIGI	TCTTTGCACI	2400
60	GTAACAGAGA	. TTCCTTTTT	CAATAATCTT	AATTCAAAGC	ATTATTAGGN	1 CTTGAAAGG	2460

A Victorian Commence

TTTGFTAATC TCCCCGTCCT TGSTAAAGGT TG

5 (2) INFORMATION FOR SEQ ID NO: 94.

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 3058 base pairs

(E) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:

1.5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:							
15	DTAAAIOOOA	AACAGACAAT	GGCATTGTCG	AAGAGCAACC	TGTTAATGAA	ATCATGTTA	60	
	AAATCAAGGT	TIGGETTCAG	CACTAAATTT	OTATEGASCT	AAGTTTATCC	TGTTTTCCAG	120	
20	PACAAACAGA	AAGTTGATCT	ATAAAADDOT	CCATTATTAG	GACCTATCAC	ACAATATCA'	180	
	TAGTTTTTT	TGTTTGTTTG	TTITTTGTTT	TTTTTTTTGG	TAAAGCCATG	CACCACAGAC	240	
25	TTTTGGGCAG	AGCTGAGAGA	CAATGGTCCT	GACATAATAA	GGATCTTTGA	TTAACCCCA	300	
25	TAAGGCATGT	CATATEATET	AAATATACTT	CTCTTTGGCT	TTTCGACATA	GAACCTCAG^	360	
	TGTTAACCAA	OATAAAEEEO	ATCAGATCTG	CAACACAGAA	ATGCTCTGCC	TGAAATTTC	420	
30	ACCATGCCTA	GGACTCACCC	ACCTATTATO	GETETTTCTG	GATCTGTTTA	ATCAATAAG	480	
	CCTATAATCA	CTTGTTAAAC	ACTGGGCTTC	ATCACCCAG3	GATAAAAACA	GAGATCATT	54(
35	THTTGGAHAT	CCTGCATCAG	CCTATTCAAA	ATTATCTCTC	TCTCTAGCTT	TCCACAAAT	600	
33	CTAAAATTO	TGTCCCAAGC	CACCCAAATT	CTCAGATCTT	TTCTGGAACA	AGGCAGAATA	660	
	TAAAATAAAAT	ATACATTTAG	TG3CTTG33C	TATGETOTOO	AAAGATCCTT	CAAAAATACA	720	
40	TCAAGCCAGC	TTCATTCACT	CACTITACTT	AGAACAGAGA	TATAAGGGCC	CACOTAGEET	780	
	DTATTTTATC	AATACCAATT	TTTCT/9GCCA	TGGCAGACAT	TGCTAATCAA	TOACAGCACT	840	
45	ATTTCCTATT	AAGCCCACTG	ATTTCTTCAC	OTOTTCOTAA	AAATTACAAT	TCCAAAGAG	900	
40	CGCCACTCAA	CAGTCAGATG	AACCCAACAG	TCAGATGAGA	GAAATGAACC	CTACTTGCTA	960	
	TOTOTATOTT	AGAAAGCAAA	AACAAACAGG	AGTITICCAGG	GAGAATGGGA	AAGCCAGGG:	1020	
50	GCATAAAAGG	TACAGTCAGG	GGAAAATAGA	TOTAGGCAGA	GTGCCTTAGT	CAGGGACCA:	1080	
	GGGCCTGAA	TCTGCAGTGC	CAACACCAAA	CTGACACATC	TCCAGGTGTA	CCTCCAACCC	1140	
55	TAGCCTTCTC	CCACAGCTGC	CTACAACAGA	GINTOCCAGO	CTTCTCAGAG	AGCTAAAACC	1200	
33	AGAAATTTCC	AGACTCATGA	AAGCAACCC	CHARTCTCTC	CCCAACCTTG	COGCATTGT	1260	
	TAATTITTAG	AACACTAGGC	TTETT CTTTC	ATHTAGTTCC	TUATAAGUAG	GRECCA GAAT	1320	
60	ADTECAGECA	CCTGCAGTGA	CATTGCTGGA	CCCCTGAAAA	CCATTCCATA	GGAGAATGG	1380	

	TTCCCCAGGC	TCACAGTGIA	GAGACATTI GA	ACADTADUES	ACTIGITTINGA	CI GCI CGCAG	144(
5	DADAAAACAG	TOCACOCACC	COATRYGACA	eurene reat	ACCEPTAGOODE	MOAGAAGM	150(
•	CAAH INGAGA	TGCTGACGTT	Gright AAS	Partertara	CAT CAGTGCA	TADDADDTAA	156(
	TCAHLACASS	AGTUATATGC	CORPTS AST	TACAAGATGT	THETTETHE	AAGCATTTTG	1620
10	ATG AATAGS	GAACTGCAAA	ngian sanga	TUTTGAAAAG	GETT CASS DAGG	ACTEGETET	1680
	AAACTGACTC	AGTGTGTCAT	CCCCGGTTAT	TTAGAATTAC	AGTTANGAAG	GAGAAACTTC	1740
15	TATAMGACTG	TATGAACAAG	PROPERTY	CATAGTGGGC	TATTACAGGC	AGGAAAATGT	1800
12	TTLAATIGE	TIACAAAACC	eamiaatacs	TOTGTCATT:	COTGLAAAAG	GCAGGAGACA	1864
	TGTGARTATG	ATCAGGAAAC	CNRCACAAAAA	CATTETTIC	AGCCCCCGTG	TIATIGTCCI	1920
20	TIMGAACIGI	THITTINIA.	TTAAAG TCAA	ATTIGTGTTG	POTTATATAT	TATTOCATOT	1980
	GIYI A BADYGA	AGCATTITCI	PTETERSACOTA	AAS AAAAAGA	a cagettetag	TAAATTATTA	2040
25	TAAAGTCGAT	GATATITCAT	CHICAGETTAT	TCTACCAAGC	TGTGCTTGTT	GGTTTTTCCC	2100
۷.,	ATGACTGTAT	TGCTTTTATA	AMI PTACAAA	TAGTTACTGA	AATGACGAGA	CCCTTGTTT	1160
	CACAGOATTA	ATAAGAACCT	DAAGAADAAD	CATATTCTGT	TGACAGICAG	CTCACAGTT	1220
30	CTTGCCTGAA	ADECTOSCITCO	CCCTCCAGTG	AGACACAAGA	TETETETT	ACCAAA STTS	1280
	AGAACAGAGC	TGGTGGATTA	ATIAAIAGIC	TICGATATCT	GGCCATGGGT	AACCTCATTG	2340
35	TARTTATOART	CAGAATGBBC	AGA GATGATO	TTGAAGTGTC	ACATACACTA	AAGTCCAAAC	:400
<i>J.</i> "	ACTATETOAG	aategegeta	AATTOCATTAA	AGAACAGGAA	ATTAATAAAA	TAAGATGATA	1460
	AGCAAATGTT	TCAG TOTAAT	GINDAACCIDAG	AAAAAA PO	AATTAATGCT	GTGTAAAATC	1520
40	GTTGAATTAG	TTTGCAAACT	ATATAAAGAC	ATATGCAGTA	TE PETENAAA	TAATGCACAT	2580
	CCTSTGGGAA	TGGAGT STTC	TAACCAATTS	CCTTTTCTTG	TTATCTGAGC	TCTCCTATAS	264
45	CTOACACTAC	AGATAATCAA	ATTAAAAGAA	TTAGAATATG	ATTATTAATA	CACTTAACAT	270
7.	CTTOTOAAAT	TAACTTICTT	CTTTCTGIGA	TAATTCAGAA	GATAGTTATG	GATICTTCAAT	276)
	GCCTCTGAGT	CATTOTTATA	AMMATCAST	TATCACTATA	CCATGCTATA	GRAGACTGGG	282
50	CAAAACCTGT	ACAATGACAA	CCCTGGAAGI	TGCTTTTTT	AA ZAAAAAA	TAAATTTCT	288
	CTCAACTAAA	TTTTTTTGG	TTGICSGITT	GTTATAAAGT	GCAACGKATT	CAAGTCCTCA	294
55	ATATIOCTGAT	CATAATACCA	TGCTATAGGA	GACTGGGCAA	AACCTGTACA	DODAADADIA .	300
دد	TGGAAGTTGC	AAAATUTUUT	TAATAATAAA	TTNTTAATCC	AAAAAAAA	MAAAAA	305

۲.

			CEC		Nº70	r :
1 - 1	INFORMATION	FUE.	SEU	-4	INC :	2 -

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1099 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: doubl€

(D) TCPCLCGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 95:

• •	1,,,,	,					
10	GGCTTYGTAG	CTGTTCCGCA	GCCCAGCCCG	GGCGCGCTCG	CAGAGICCTA	GECHETGCGF	€(
	GGCNTCCTGI	CTCCTCCCTC	ರ್ಣಾಚಾರಚರಾರ	GCGGCCCGCG	CCTCCGCGGT	GCCTGCCTTC	120
15	GITCTCAGGI	TGAGGAGCTC	AAGOTTGGGA	AAATGETGTG	CATTCCTIGI	ATCGTCATTC	180
	CAGTTCTGCT	CTPGATCTAG	AAAAAATTCC	TGGAGCCATA	TATATACCT	CIGGTTTCCC	240
20	CCTTCGTTAG	TOSTATATGS	CCI AAGAAAG	CAATACAAGA	ATCCAATGAT	ACAAACAAAG	30:0
20	GCAAAGTAAA	CTTTAAG3G1	GCABACATGA	AT GEATTACC	AACAAAAGGA	CCAACAGAAA	36(
	TCTGTGATAA	AAAGAAAGAC	TAAAGAAATT	TTCCTAAAGG	ACCCATCAT	TTAAAAAATG	410
25	GACCTGATAA	TATGAAGCAT	CTTCCTTGTA	ATTGICTCTG	ACCTTTTAT	CTGAGACCGG	48(
	AATTCAGGAT	AGGAGTCTAG	ATATTTACCT	GATACTAATC	AGGAAATATA	TGATATCCCT	54(
30	STAAAAITTA	TAGTTAGTTA	TATTTAATGA	CCTCATTCCT	AAGTTCCTTT	TTCGTTAATG	600
30	TAGCTTTCAT	TTCTGTTATT	GCTETTTGAA	TAATATGATT	AAATAGAAGG	TTTGTGCCAG	660
	TAGACATTAT	GTIACTAAAT	CAGGACTTTA	AAATCTTTGG	TTCTCTAATT	CATATGAATT	720
35	TGCTGTTTGC	TCTAATTTCT	TIGGGGTCTT	CTAATTTGAG	TGGAGTACAA	TTTTGTTGT	780
	AAACAGTCCA	GTGAAACTGT	GCAGGGAAAT	GAAGGTAGAA	TTTTGGGAGG	TAATAATGAT	840
40	GTGAAACATA	AAGATTTAAT	AATTACTGTC	CAACACAGTG	GAGCAGCTTG	TCCACAAATA	900
40	TAGTAATTAC	TATTTATTGC	TCTAAGGAAG	AAAAAATTA	GATAGGGAAA	AGGGGGAAAC	960
	TTCTTTGAAA	AATGAAACAT	CIGTIACATT	AATGTCTAAT	TATAAAATTT	CATTOOTAAT	1020
45	TGCATTTCTT	CTGTTCCTAC	AAATSTATTA	AACATTCAGT	TTAACTGGTA	AAAAAAAA	1080
	AAAAAAACCC	GGGGGGGGG					1099

50

55

(2) INFOFMATION FOR SEQ ID NO: 96:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1580 base pairs

(B) TYPE: nucleic ació

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:

ESTIMATION STATISTICA CATAMENTA OTTIMARA CARRESPAT SCANDAMINA BULTURIBRE RIPATAMA ACCUMENTA OTTIMINADA CARRESPATA CONTURBANI ACCUMENTA ACCUMENTA ACCUMENTA ACCUMENTA CONTURBANI CONTURBANIA ACCUMENTA ACCUMENTA CONTURBANIA CONTURBANIA ACCUMENTA ACCUMENTA ACCUMENTA CONTURBANIA CONTURBANIA ACCUMENTA ACCUMENTA ACCUMENTA CONTURBANIA CONTURBANIA ACCUMENTA CONTURBANIA ACCUMENTA CONTURBANIA CONTURBANIA ACCUMENTA CONTURBANIA ACCUMENTA CONTURBANIA CONTURBANIA ACCUMENTA CONTURBANIA ACCUMENTA CONTURBANIA CONTURBANIA ACCUMENTA CONTURBANTI CONTURBANIA ACCUMENTA CONTURBANIA ACCUMENTA TOTOCOCATO CONTURBANIA ACCUMENTA CONTURBANIA CONTURBANIA ACCUMENTA CONTURBANTI CONTURBANIA ACCUMENTA CONTURBANIA CONTURBANIA ACCUMENTA CONTURBANI CONTURBANIA ACCUMENTA CONTURBANIA CONTURBANIA CONTURBANIA CONTURBANIA ACCUMENTA CONTURBANIA CONTURBANIA CONTURBANIA CONTURBANIA CARRESPANTA CONTURBANIA CONTURBANIA CONTURBANIA CONTURBANIA CARRESPANTA CONTURBANIA CONTURBANIA CONTURBANIA CONTURBANIA CONTURBANIA CARRESPANTA CARRESPANTA CONTURBANIA CONTUR		GECAGAGAM.	14GZZZZZZZZZ	TTCATGAAAA	AANGCAGTCC	CTTAACTI CA	GTTTGACARA	ŧ
######################################	4	GTGGAGFTS CC	percondicionar	TOACCACAC	GATICTOCTO	Mindi-digi	GOTTTIGAGGG	2:
10 ACCRECISA ACCRECATE TETTE SECON PERCENTED: TEACH TOSS PLANFORMS: 1 TOCKNOWN ASSOCIATION TETTE SECON PERCENTS. 1 TOCKNOWN ASSOCIATION AS	٠	COLACINIAN	CTCATTTICA	CATCAGTGAT	GTTT(#C-3AC)	CAGGU POACT	CCATCTGJAC	18
TOTOROSTOR HACURAGNIA ENGUAGEOU GENERATA STORT THAN GUNCOUSA. OGUNTOSTO ACTURIANA MATERIALA CONTRANA TOTTORIAN GUNCOUSA. AMUSTORAA TUSAAAANU AMARKAACIA COSTAAAA TOTTORIANS TOTORASTAA. AMUSTORAA TUSAAAANU AMARKAACIA COSTAAAA TOTTORIANS TOTORASTAA. AMUSTORAA TUSAAAANU AMARKAACIA COSTAAAA TOTTORIANS TOTORASTAA. AMUSTORAA TUSAAAANU AMARKAACIA COSTAAAA TOTTORIANS TAMAACIATU 20 OTTORASTITI COURTISTUS TORGITTISSU ADMINISTIC TITORAAANIS ACTURITISSU COURTISSITTI COURTISTUS TORGITTISSU ADMINISTIC TITORAAANIS ACTURITISSU COURTISSITTI COURTISTUS TORGITTISSU ADMINISTIC TITORAAANIS ACTURITISSU COURTISSITTI COURTISTUS AMARKAACIA GARVITISTI TUSAAAANIS ASTITORIAGA 25 OOTTISSITTI COURTISSU AMARKAACIA GARVITISTI COURTISAGA GUNGOARIA GUNCAATISSU AMARKAACIA GARVITISTI TUSAACAACIA COURTISAGA COURTISSITTI COURTISSU AMARKAATIS COURTISTI ATTITATISSI CITOROSTIGI COURTISSITTI COURTISSU AMARKAATIS COURTISTIC ATTITATISSI CITOROSTIGI COURTISSITTI ATTITORISSU AMARKAATIS COURTISCOTTIGI ATTITATISSI CITOROSTIGI TOVAAAACISI TOTUVISSISE ATTICISTICIA CORCCOACIA CARTITORIS CANACTACCII 40 TITOCOTTIGI TITURICITISTI TOCIOTOTICA CORCCOCATI TOCCOTTICTI COURAGAACI ACTITORICI TITURICITISTI TOCIOTOTICA CORCCOCATI TOCCOTTICTI COURAGAACI ACTITORICI TOCOTTITIC ACTITICACTIS CANTIGGERA CARTITORI CITORAGAA COCCACCATA ACTITICISA ATTICISACTIS CANTIGGERA CARTICOCTI COURAGAGI COURCCATA ACTITICA COURTISTICI TATOLATACII TITACOCTICA CATTITORIAGI COURCCATA ACTITICACII ACCIGITICI TATOLATACII TITACOCTICA COURTICATI TOTTORICITA GUNUTAGACO ATTICOCTICI TATOLATACII TITACOCTICA COURTICACII ACCIGITICA TATOLATACII TITACOCTICA COURTICACII ACCIGITICA TATOLATACII TITACOCTICA COURTICACII ACCIGITICA TATOLOCTICA COURTICACII TITACOCTICA COURTICACATII TITACOCTICA COURTICACATII TITACOCTICA COURTICACATII TITACOCTICA CARTICOCTICA COURTICACATII TITACOCTICA COURTICACATII TITACOCTICA TITACOCTICA COURTICACATII TITACOCTICA COURTICACATII TITACOCTICA TITACOCTICA COURTICACATII TITACOCTICA COURTICACATII TITACOCTICA COURTICACATII TITACOCTICA COURTICACATII TITACOCTICA COURTICACA		AGATGAGAGG	iksaat agaac	AASTNAAAAA	GGAAGAGAGA	ACASC PCACA	AAAAAAACAAA.	24
CONTINUES CONTINUES AND ANALOGO CHINOCATTA COSTOCICATE OSTOCICATE ANTOSTIANA TOCARANAMA ANALOGA CONTINUANA TOTTICIONO COSTOCICATE ANTOSTIANA TOCARANAMA ANALOGA COSTOCICADA TOTTICIONO COSTOCICATEMA ANTOSCICARO CATURCASAMA ANALOGACA CONTINUESTO TOTTICIAMATE ANALOGATE CONTINUESTI COCICIONES TOTICICIONES ATRACTICATE TOTTICAMATE ACTICIONES COTTICIONES CATURCAS ACCUMUNICAS ANALOGAMEN CATURCATA COCCICIONES COTTICIONES CATURCAS ACCUMUNICAS ANALOGAMEN CATURCAS COCCACOSE CONTINUESACA COTTICIONES CATURCAS ANACOMINA ANALOGAMENA CATURCAS COCCACOSE CONTINUESACA COTTICIONES CANALOGOS ANACOMINA COCCACOSE CONTINUESACA COTTICIONES CANALOGOS ANACOMINA COCCACOSE CONTINUESACA COTTICIONES CONTINUESCACIO CANACIDANA ANALOGAMENA ANALOGAMENA ANALOGAMENA CONTINUESACA TOTARANCONE CIULINACANI CANAGARRAMA ANALOGAMENA CONTINUESACA CANACITACOS ANTICOTTURI TELEVICISCI TUCCIDIOTCA CONCOCAMIC TOCCONTENTI COCCAGGAME ANTICOTTURI ANGLERICIA ANTICIGACO CANACITACO INCOCCACIO COCCACIO COCCACI	10	DAAEMAAHDA	ATGAAAGTCG	TTTTTIGGGCA	CCCCIMEDE:	nd AGN tikyag	CUAGGGGGTT	30
AMONITORIA TOGRADANI. ANDMONITO COTTAINA TOTTOTIMO TOTTAINTM ANTOGOTORIA COMPANIANI ANAMANANI COTTAINA TOTTOTIMO TOTTAINTM ANTOGOTORIA COMPANIANI ANAMANANI COTTAINA TOTTOTIMO TOTTAINTM ANAMANANI 20 COTTOGRATTI COMPTINI A COGNINCTRI TOTTOTITOTI TOTTAINAMO ATTOTTOTORI COMPTINITI COMPTINI TOTTOTIMI ANAMANANI ACTOTTOTORI COMPTINI COMPTINI ANAMANANI GARROTORIA TOTTOTIANI ANAMANANI COMPTINI COMPTINI ANAMANANI GARROTORIA TOTTOTIANI ACTOTTORIA COMPTINI GARROTORIA ANAGOTARIA GARROTORIA COMPTINI COMPTINI COMPTINI COMPTINI ANAMANANI GARROTORIA GARROTORIA COMPTINI COMPTINI COMPTINI COMPTINI ANAMANANI GARROTARIA GARROTORIA COMPTINI COMPTINI COMPTINI COMPTINI COMPTINI COMPTINI COMPTINI COMPTINI COMPTINI TOTTOTIMI COMPTINI CANTARIA ANAMATINI ARAMANANI COMPTINI COMPTINI COMPTINI TOTTOTTIMI COMPTINI CANTARIANI CANTARIANI ANAMATINI COMPTINI COMPTINI COMPTINI ANAMANI CANTARIANI CANTARIANI CANTARIANI ACCOMPTINI COMPTINI COMPTINI ANAMANI CANTARIANI CANTARIANI CANTARIANI COMPTINI COMPTINI COMPTINI ANAMANI CANTARIANI CANTARIANI CANTARIANI CANTARIANI COMPTINI COMPTINI ANAMANI CANTARIANI CANTARIANI COMPTINI COMPTINI COMPTINI COMPTINI ANAMANI CANTARIANI CANTARIANI CANTARIANI COMPTINI COMPTINI COMPTINI COMPTINI COMPTINI CANTARIANI CANTARIANI CANTARIANI COMPTINI COMPTINI COMPTINI COMPTINI COMPTINI CANTARIANI CANTARIANI COMPTINI COMPTINI COMPTINI COMPTINI COMPTINI CANTARIANI CANTARIANI COMPTINI CANTARIA COMPTINI COMPTINI COMPTINI COMPTINI COMPTINI CANTARIANI COMPTINI COMPTINI CANTARIA COMPTINI COMPTINI COMPTINI COMPTINI COMPTINI CANTARIA COMPTINI CANTARIA COMPTINI CANTARIO COMPTINI COMPTI		TGCCACGCCA	GACCAAG 43A	ACCAGACCC	GIACCAGIAI	STGST TTUAA	GGACCCGGAL	3 t
ARTSSTEARA TOSARARAS. RAMINARIA CONTARA TOTTTOTAS DECEMBERA ARTSSTEARA CATANDASA. RAMINARIA STOCKARA TOTATOTAS DECEMBERA 20 COTTOSATT COUNTS ON A COGNISCISC CUTOSTITO TOTARAMO ACTIONOGO CONTISST COUNTS TOTTO SO A COGNISCISC CUTOSTITO TOTARAMO ACTIONOGO CONTISST CONTOGRES TOTSTORS. REGISTRATA CAGAS PORS TOTATOGOGO TOTOGREST CONTOGRES REGISTRADA SAGONARA GRONTOGO TOTARAMA ACTIONARA 25 CONSTIGUES GRONKARIAS ARASONARA GRONTOGO TOTARAMA ACTIONARA COTGRESARS GRANKARIAS ARASONARA GRONTOGO CICCOACTO CORRESPADA COTGRESARS GRANKARIAS ARASONARAS CONGRESS CONTRACTORS CORRESPACA COSTITUCIO GUARACTOS ARASONATOS CONGRESSOS CONTRACTORS CORRESPADA COTGRESARS CONSTITUTOS ARASONATOS CONGRESSOS CONTRACTORS CONTRACTORS CONGRESARS CONTRACAS ARASONARIOS CONGRESSOS ACTUATORS CONTRACTORS CONGRESARS CONTRACAS ARASONARIOS CONGRESSOS CONTRACTORS CONGRESARS CONTRACAS CONGRESSOS ARASONARIOS CONTRACTORS ACUAGORATO RESPONSE ENTRACTORS GRANKENES TRASSARSAC GUATUTOCRA 135 AURAGORATO RESPONSE ENTRACTORS GRANKENES TRASSARSAC GUATUTOCRA 140 CONGRESARS CONTRACAS AURAGORDA ARABITICA CARTURAS CARROTAGOT 141 AURICANTA ARASONARIO CONTRACTOR CONCRETE COCCAGONA 151 CONGRESARS ARASONARIO RECURSIONI CARROTAGO CONCRETE CONCREGAR 164 CONCRECATOR ARASONARIO RECURSIONI CARROTAGO CARROCACO ARASONARIO 165 CONCRECATOR GRANKENCO RECURSIONI CARROCACO CARROCACO ARACONACIO ARACONACIO CONCREGARAS 165 CONCRECATOR ARASONARIO RECURSIONI CARROCACA CARROCACACO ARACONACIO ARACONACIO CONCREGARAS 165 CONCRECATOR ARACONATICO RECURSIONI CARROCACA CONCREGAR ARACONACIO ARACONACIO ARACONACIO CONCREGARAS 1713TITUDA ARACONALICO RECURSIONI CONCRETITI CARROCACA CONTICCOLO CONCREGATORIO 1713TITUDA ARACONALICO RECURSIONI CONCRETITI GRAGRAGORIA CONTICCOLO CONCREGAR TOTTUTUDA ARACONALICA CARROCACIONI CARROCACA CONTICCOLO CONCRETICO CONCRETICO CONCRECATORA CONCRETICO RECURSIONI CONCRETICA CONCRE	15	CGRIATGGTC	ACCICAGACAC	OACACCTOAA	CACAGCACTA	COSTO CATO	CGTTCTCATG	4:
CONTROGATI CIVITARE SA COGARROTES TUTUTITADO TUTURARANTO AUTUTOTOS CONTRESTITI CIVITARE SA COGARROTES TUTUTITADO TUTURARANTO AUTUTOTOS TREGOCCATO CUACIDANES TUTUTITASES ADADUTADAS DAGRERADO TREGOCOACIDE CUARRAGEA 25 COTREGOS GATRIATESA ARAGONARSA GARROTESA DESACRICACIDES CUARRAGEA COTREGOS GATRIATESA ARAGONARSA GARROTESA DESACRICATA CONTINAGO COTRAGARAS GUARRACEAS ARAGOSACOTI GROCAGARA DESACRICATA CONTINAGO COTRAGARAS GUARRACEAS ARAGOSACOTI GROCAGARA DESACRICATO CARROTESACA COTRAGARAS GUARRACEAS ARAGOSACOTI GROCAGARA DESACRICATO CARROTESACA COTRAGARATEAN TUTURARESES ARAGOSACA ARAGONARA ARENTINATO PRESIDINATO TRATUTOTESA ST. ATURARACEANTI AURITECCTO I SATURITIES GARRACEAS ARENTICARS CARROTECAS TUTURARACEA COMUNICAMO CARRAGORAA ARABONIRA CARTONARS CARROTECAS TUTURARACEA COMUNICAMO CARRAGORAA ARABONIRA CARTONARS CARROTECAS 40 ATURICATURA AGUARTICESA ANTONARA CARROCACA CARCONARA CARROTECAS COCCACCATOR TOCOCOTI DE AUTURIATARA CONCARACEA SATURACETTI CUARLAGOTA 45 ACOCCACCATOR TOCOCOTI DE AUTURIATARA CONCARACEA SATURACETTI CUARLAGOTA COCCACCATOR TOCOCOTI DE AUTURIATARA CONCARACOR SATURACECOTI CONCACAGONA CONCACCATOR ARESOCACION ARCONARIO CARROCACA SATURACECOTA CARCONARA COTRAGARIA ARROGACION ALCONARIOS CARROCACA SATURACECOTA ARGONETTO COTRAGARIA CORROCALION ALCONARIOS CARROCACA SATURACECA ARGONETOS COTRAGARIA GRACOLION ALCONARIOS CARROCACA SARURACACO ARAGONETA COTRAGARIA GRACOLION ALCONARIOS CARROCACA SARURACACO ARAGONETA COTRAGARIA GRACOLION ALCONARIOS CARROCACA SARURACACO ARAGONETA COTRAGARIA GRACOLION ALCONARIOS CARROCACA CONTOCCACA CONTOCCACACA CONTOCCACACA CONTOCCACACACA CONTOCCACACACACACACACACACACACACACACACACACAC	1.	AAADTTEETIAA	DOGANANS.	ARARCAACTA	COCOMIAAAA 1	dental services	DOD DAAGTAA	4+
CONTRACTO CONTRACTOR TOTAL TOT		AAT GGTTGAG	CATTGCAGAS	AAAAAAAA	CACADOOC TO	and Yad andat	AAAAAMCATKU	54
TOTOGOCCATO CTACTARTOS ACTURASACAS GROSTOSTOS TESSATIAGRAS ACTURASACA COTOGOGRAS GROSCATORA ACTURAS ACTURAS ACTURAS COTOCCACOST CORRESPONDO COTOGOGRAS GROSCATORA ACTURADA COTOCCACOST CORRESPONDO COTOGOGRAS GROSCATORA ACTURADA COTOCCACOST CORRESPONDO COSTITUCIO GOTRACTORA ARAGERATOS COTOCCACOST CARROSPACA COSTITUCIO GOTRACTORA ARAGERATOS COTOCCATOROS CARROSPACA COTOGOGRAS TOTOGOGRAS ACTURATORA ARAGERATAS ACTURATORAS COTOCCACACA COTOGOGRAS TOTOGOGRAS ACTURATORAS ARAGERARAS ACTURATORAS COTOCCACACACACACACACACACACACACACACACACAC	20	CIMICGAMIT	CERNALASS BY	COGAWGCTGO	TOTALITATION	OTAAAACCTT	ACTICIPATES	€11
CONSCIONS GAT VOLTERS ANAGORASSA GRADOCTISA COCCCACION CLASSAGRICA GERROPRO COCCENTRAS ALATAMACO TEMPOGAMSA TRANSISCITTA COCCETTAMOS 30 COTGASAMOS GENERACIONE ANAGORATIS COTGACAGGA COTCATORIO CANGORANA COSTITUCO I GULANCIONE ANAGORATIS COTGACAGGA COTCATORIO CANGORANA GENERATURA TETRA GRADOCTO ANAGORATIS COTGACAGGA ACCITATE TRANSISCITATI ST ATUMACAMONI TETRA GRADOCTO E SATUNDOS ANAGORANA ANAMITICA CANGORANA GENERATORIO I TOURAMACONO CIUNALACANI CANGAGGAMA ANAMITICA CANGORANA CANGORANA 1 TETROCTORIO TETRACOCTRA COCCUCATA TOCCOCCAGORANA 1 ATURICANTIA AGORASIGISMA ATUCISACOS CANGORANA CANGORANA CANGORANA 1 COCCACCATOR TOCCOCTUCIO ACCIGISTO CANGORANA CANGORANA 1 ACCOCCACICA TOCCOCTUCIO ACCIGISTO TETROTATALI TUNACCEGGO AGACTICATO 1 ACCOCCACICA ANTORIO ACCIGISTO TETROTATALI TUNACCEGGO AGACTICATA 1 ACCOCCACICA CANGORANA CANGORANA CANGORANA 1 COCCACCATOR GENERALGACO ATURISTICA SECUNOCICA CANGORANA 1 COTGARCATAL GENERALGO ATURISTICA SECUNOCICA CANGORANA 1 COTGARCATAL GENERALGO ATURISTICA CANGORANA CANGOCTAGO COTTUCANA 1 COTGARCATAL GENERALGO ATURISTICA CANGORANA CANGOCTAGO 1 COTGARCATAL GENERALGO ATURISTICA CANGOCTAR CANGOCTAGO COTTUCANA 1 COTTUCANA CANGOCTAGO ATURISTICA CANGOCTAGO COTTUCANA CANGOCTAGO COTTUCANA 1 COTTUCANA CANGOCTAGO ATURISTICA CANGOCTAGO COTTUCANA C		CCTTTGGTT	CTCTCTGTTG	TOTOTOTO	ATHARTTARFA	TAGAG (CTG)	indulăniêl@	61
CONSCISSION GENERALISM ARRECTANCE GRANCETTA COCCACIST CLASSAGACA GENERALISM COCCGAINS ACAIAAAACC TOACCGAASA TERRITECTTA COCCETTSAGE 30 COTGABAAGO GERBEACONS ARRESACTI GRANCAGGA COTGATORCO GAAGDREACA COSTITUCCO GENERACTOR ARRESACTI COTGACTOR ACTIVATERS CITISGOSTGT GENERALISMO TITIS 1933 ACTURITIES GARGAGERAAS AGRICITATI TRIATTOTGA 1 ATCAGCAATT ACAITOCOTG ERACTETIS GARGAGERAAS AGRICITATI TRIATTOCAGE TOVARAATISC CIMALACAAT CARGAGRAAA ARRAITECA GARTICAGO GARGAGAAGA TITICOTTIST TICATCISSI TOCCICIOCOTA CORCOCCATA TOCCICIOTT COCCAGOAAG 1 ACTICACACACA AGRICICA ATCOCCATOR GARTAGGEAC GACATACAGT COCACCACACA TOCCCCTOIC ACTIVIATAAA COTGAAAGTS BATCACCTT COGATAGTTA 45 ACCCCCACACA ACTIVIAGACC ACAGOCCAS GARTICAGA AGRICACAGA COCTGATORT GENERALGACC ACAGOCCAS GACACACACACAC COTTCCAAAG 1 COCTGATORT GENERALGACC ACAGOCCAS GARTICAGAC GACACACACAC COTTCCAAAG COCACCACACA ACTIVIAGACC ACAGOCCAS GARTICAGAC GACACACAC COTTCCAAAG COCACCACACACA GACCACACA COC ACTIVIAGACAG GARTICAGAC GACACACAC AGAGCTAGG COCACCACACACAC GACCACACA GACCACACAC GACACACAC	25	THIGHCOATI	CTACTAA JIG	MODAGASITSA	GACGCCCCC	TAADHTAGET	AGTITGGACA	7:
CONTRACTOR OF ACCOUNTS CANTOCAGE CONTRACTOR CANTOCAGE CONTRACTOR OF ACCOUNTS CANTOCAGE CONTRACTOR CANTOCAGE CONTRACTOR OF ACCOUNTS CANTOCAGE CONTRACTOR CONTRACTOR CONTRACTOR CONTRACTOR OF ACCOUNTS CANTOCAGE CANTOCAGE CANTOCAGE TOVARABONE CONTRACANT CARGAGERAA ARANTIACA CANTOCAGE CANCOTAGE TOVARABONE CONTRACTOR CONTRACTOR CANTOCAGE CANCOTAGE TOCCOTAGE TOVARACTOR ACCOUNTS CANTAGERA CACATACAGE CONTRACTOR TOCCOTOR ACCOUNTS CANTAGERA CACATACAGE CONTRACTOR CONTRACTOR ACCOUNTS CANTOCACA CACATACAGE CONTRACTOR CONTRACTOR ACCOUNTS CANTOCACA CACATACAGE CONTRACTOR CONTRACTOR ACCOUNTS CANTOCACA CACATACACT ARAGERTAGE TOTAGERALA CONTRACTOR ACCOUNTS CANTOCACA CACATACACT ARAGERTAGE CONTRACTOR CONTRACTOR ACCOUNTS CANTOCACA CACATACACT ARAGERTAGE CONTRACTOR CONTRACTOR ACCOUNTS CANTOCACA CACATACACT ARAGERTAGE TOTAGERALA CACATACACT ACCOUNTS CANTOCACAC CONTRACTOR ACCOUNTS CONTRACTOR CACATACACT CONTRACTOR CONTRACTOR CONTRACTOR CONTRACTOR CONTRACTOR CONTRACTOR CACATACACT CONTRACTOR CACATACACT CONTRACTOR CONTRACTOR CACATACACT CONTRACTOR CONTRACTOR CONTRACTOR CACATACACT CONTRACTOR CONTRACTOR CACATACACT CONTRACTOR CONTRACTOR CACATACACT		CCTGGTGG 73	GATHCATGRE	AAAGTCAGGA	CARTOCCEED	CICCCACAG	CCAGGAGGCA	71
CGETTISCII GITAACTOSC AFAGCARTIG CONGOTTOT ACTIMATESS CINSGETGT GIAGRAMICAI TITU-19555 ACTGESSAGA ARGAIGARAS ACEICTIATI TETATTOTGA ATCAGCARTI ANATHOCOUG ISATIATIES GARGAGTETE TAGGARAGAC GITTITOCAGE TIVARARANSC CIMENACANI CARGAGGARA ARAANITACA GRATTICAGE CARGOTACGI 40 TITICOTTIGE TILAVOTECT DECTOTOCA CORCOCCATO TOCCOUTOTT COCCAGORAG ATGETCARTIA AGGASTICISA ATTOTSACTE CARTAGGCAG JAGTGOCCAR CACATACAGE 45 COCACCATOR TOCCOUTOTC ATTIVATARA CONCARAGES SATICACTITI CINGATAGTA ACCOCCCATOR ANTERGACO ACCIGISTOT TATCIATATI TIRACCEGGG AGACTGTTGT 1 COCCAGORIA ANTERGACO ANACIGISTOT TATCIATATI TIRACCEGGG AGACTGTTGT 50 TINGACATAT GESCORTOC ATTIS SECOND GRATTICACA SERCACACCI ARGCTGTGG CMAYTOCOGG AGARANTASAT TITCONTITI GROGROGGCA CUTTOCCTGT TOTTCAGTTC TITSTITUBAA GGGROGGING		@#990999911 [.]	00000377393	COAAAAACCA	TOATOGAAGA	TERRET SCTTA	COCOTTSAGE	64
GIAGAATGAT TYPE-199990 ASTERBAGA AAGATGAAAS ADETETTATT TETATTOTGA ATCAGCAATT ATATTCCCTG TRATTATTES GAAGAGTES TAGGAAAGAC GITTITCCAG TYVAAAAATGC CYMALACAAT CAAGAGGAAA AAAAATTACA CAATTICAGG CAAGCTACGT 40 TITCCTTTGT TYPE-199900 TCCCTCTCTCA CCACCCCATC TCCCTCTCTT CCCCAGGAAS ATGTCAATTA AGTASTGUGA ATTCTGACTS CAATAGGCAC PACTGCCTAA CACATACAGC cccaccatca Tocccttcic Atternataaa cctcaaaagts Sattcacttt Cigatagtta 4.5 ACCCCCAIAA ATGTGCAC ACCTGTGTC TATCTATATT TOAACCEGGG AGACTGTTGT 1 CCTGGGCATG GGGAATGACC ATGGCTGS GCTIACCTCA CASCCCCTAC CCTTTCAAAAG 50 TINGACATAT GGGCCATCCC ATGGCACA GAATTCCACA GGACCACCT AAGGCTGTGG GMAYTGCGGG ACAAATAGAT TITCCATTTI GAGGAGGGCA CTTTCCCTGT TGTTCAGTTC 1 TTSTTTTSAA GGGAGGTNG	30	CCTGAGAAGG	GCAGGATCAG	AAGGGACCTI	GGCACAGCGA	CCCOTADICO	CAAGTGGACA	90
ATMAGRANT ANATHORNO I SATIATIFS SAAGASTSTS TABSAAASAN STITTOMS 1 THUAAAANSK CHUMARAAN CAAGASSAAA AAAANTARA GAATTIGASS GAAGCTACGI 1 40 TITOOTITIST THUATCISCT TECHTOTOA GRACGCCATE TOCCTUTOTI COGCAGNAS 1 ATMICAATIA AGGASTISUSA ATTOUGACTS GAATAGGEAC JAGTGCCCAA GAGATAGAGC 1 COGCACCATCA TOCCCTTOTIC ATMICATAAA COUGAAAGTS SATTGCCTTT CIGATAGTTA 1 ACCCCCCATAA ATSINGAACH ACCIGTSTCT TATOTATATI TUAACCEGGS AGACTGTTGT 1 COTGAGCATG GAAGATGACC ATMICCCAG GAATTOCACA GASTCCCCAG COTTOGAAAG 1 50 TINGACATAT GOSCCATCCC ATMICCCAG GAATTOCACA GASTCCCTGT TOTTGCAGTTG 1 CMAYTGCGGS AGAAATAGAT TITCCCATTT GAGGAGGGCA CTTTCCCCTGT TOTTGAGTTG 1 TITGTTTIBAA GGGAGGTING 1		C032111 3C01	GTTAACTOGO	AAAGCAATTG	carecarrer	ACTITIAN 333	CITGGGGTGT	95
ATMAGRATT ANAMOROUS ISATIATINS SAAGASTSTS TAGGAAAGAC GITTITOOAG 1 TIMBAAACING CIMBIACANI CAAGAGGAAA AAAANTIACA GAATTICAGG CAAGCTACGI 1 40 TITICOTTIGI TIMBICINGI TICCICITOOA GCAGGCCATE TOCCICITOTI CCCCAGGAAG 1 ATMICAATIA AGGASTGIGA ATMCIGACTIS CAATAGGCAC MAGRICOCAA CACATACAGC 1 CCCACCATCA TOCCCITICIC ATMILATAAA CCIGAAAGTIG GATTCACTIT CIGATAGTTA 1 ACCCCCATAA ATMIGACACH ACCIGIGICI TATCIATATI TIVAACCEGGG AGACTGTTGT 1 CCTGGGCATG GGASATGACC ATMIGCTGS GCTIACCTCA CASTCCCCGAC CCTTTCCAAAG 1 TIMGACATAT GGGCCATCC ATMIGCCAG GAATTCCACA MAGCACACT AAGGCTGTGG 3 CMAYTGCGGG ACAAATAGAT TITCCATTIT GAGGAGGGCA CTTTCCCCTGT TOTTCAAGTTC 1 TITCITTIGAA GGGAGGTIGC 1	35	GIAGAATGAI	364501-70000	AGTG FGGAGA	AAGAT GAAAG	AGGICMA M	TRIATMOTGA	101:
40 TITICITIES TELETITISC TECTETECA COACCCATE TOCCTITETT CECCAGEAGS 1 ATSTOLATIA AGGASTGERA ATTOCATTE CALAGRAC LAGROCCIA CACATACAGC 1 COCACCATCA TOCCCTTUIC ATTICATALA COMBANGES SATIVACITY CIGALAGITA 1 ACCCCCATA ATSTSCANCE ACCIGIGICE TATOLATALE TELACOROGIA AGACTISTICE 1 COTOGRACAS GRASALGACO ATMALISTICS COLLACITA CASOCCCCAC COTTUCAAAG 1 CMAYNOCOGIA ACAANTAGAT TITICCATTI GAGGAGGGCA CTTICCCTGT TOTTCAGTTC 1 TITITITISAA GGGAGGING 1		ATCAGCAATT	ATATTOOTTG	I GATTATITG	GAAGAGTGTG	TAGGAAAGAC	GTTTTTCCAG	108
AMSTORATIA AGGASTGUSA ATTOUGACTS CRATAGGUA JAGTGUCCAA CACATACAGT 1 COCACCATOR TOCCOTTUTO ETETLATARA COMBRAGETS SATTOROTTT CUGATAGETA 1 ACCOCCATOR ACCEPTATE ACCEPTATE TATOURANT TORROTCOGA AGACTGETGT 1 COTEGRACIES GENERAGES ATMAIGNING COTTEGRARAGE 1 THINGACATAT GESCOCATOR AMERICAN GRATTOCACA CHACACACT RACCOTTUGARAGE 1 CMAYURGOGG ACARATAGAT TITOCATTI GAGGAGGGCA CUTTCCCTGT TOTTCAGTTC 1 THETTITISAA GGGAGGING 1		TTCAAAAATGC	TAASA PA PTS	CAAGAGGAAA	AAAAATTACA	CAATTICAGG	CAACCTACGT	104
45 ACCCCATCA TOCCCTTCTC ATTIVATAAA CCTCAAAGTS SATTCACTTT CIGATAGTTA 1 ACCCCCATAA ATSIGCACCI ACCIGTGTCT TATCTATATI TVAACCEGGG AGACTGTTGT 1 CCTGGGCATA ATSIGCACCI ATVATGCTGS CCTTACCACA CASTCCCCTAC CCTTTCAAAG 1 50 TINGACATAT GGGCCATCCC ATVATGCCAG GAATTCCACA CASTCCCCTGT TCTTCAGTTC 1 CMAYTOGGGG ACAAATASAT TITCCATTTI GAGGAGGGCA CTTTCCCTGT TCTTCAGTTC 1 TTGTTTTGAA GGGAGGTNG 1	40	TTTCCTTTGT	TTIATITIET	TOCTOTOTOA	CTADUODADO	TOCOCUCTOTT	CCCCAGCAAG	120
45 ACCCCCATAA ANGIGCACGI ACCIGIGICI TATCIATATI TIAACCEGGG AGACIGITGI 1 CCTGGGCATG GGAGATGACC AINAIGCIGG GCTIACCITCA CAGCCCCAC CCTITCCAAAG 1 50 TINGACATAT GGGCCATCCC AINGGCCAG GAATTCCACA GGACACCT AACGCTGTGG 1 CMAYCGCGGG ACAAATAGAT TITCCATTII GAGGAGGGCA CTITCCCTGT TCTTCAGTTC 1 TIGITITIGAA GGGAGGING 1		ATTRACTERTA	AGTARTGIRA	ATT CTGACTS	CAATAGGCAC	JAGDOPODAA	CACATACAGE	126
COTEGRADATE GENERATERACO RIMARIGOTER MODIROCTOR CARROCCIRA COTTUCARAGE : 50 TINGRORIAT GENERATERACO RIMERMOCRE GRATICORCA PERCACRACOT RACCOTETO : 1 CMAYURGOUS ROBRATASAT TITUCATUTI GROGROGICA CUTUCCOTET TOTUCAGUTO : 1 TUSTUTURAR GENERATERA	45	AGTACCACCO	TOCCCCTTOIC	ACYTOTI ATAAA	CCUCAAAGTG	3ATTCACTTT	CIGATAGTTA	13.
50 TINGACATAT GOGCCATOCC ATMINISCOAG GAATTOCACA MARCACACOT AAGGCIGTGS 1 GMAYTOCGG ACAANTASAT TITCCATTTI GAGGAGGGCA CTTTCCCTGT TOTTCAGTTC 1 TITGTTTTGAA GGGAGGING 1		AA TAGGGGGGA	ATGTGCACGT	ACCIGIGICI	TATCTATATI	TTAACCEGGG	AGACIGTIGT	138
CMAYTYGGGGG ACAANTASAT TITYCCATTII GAGGAGGGCA CTTTCCCTGT TGTTCAGTTC 1 TTGTTTTGAA GGGAGGING 1		CCTGGGGCAG	GENERAT GACC	AT-ATGCTGG	ASTIACCICA	CARDODORAC	CCTTTCAAAG	145
TTSTTTTSAA GGGAGETNG	5()	TA KACADITT	000 "A 000669	ATM & BOCCAG	GAATTCCACA	TODACACAFF	AAGGCTGTGG	36.
115111511 COCIN CINC		GMAYTGGGGG	ACAAATAGAT	TTTCCATTI	GAGGAGGGCA	CTTTCCCTGT	TGTTCAGTTC	156
	55	TTSTTTTSAA	GGGACGTNG					158

(2) INFIRMATION FOR SEC II NO: 97:

	(i) SEQUENCE CHAFACTERISTICS: (A) LENGTH: 678 base pairs	
.	(E) TYFE: nucleir acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97	
1.6	POAGTAATOT DIAITOGASED AEDAGTTASD AUATAGAGSS TETAAITEGA INTITUTUADA	€ -
10	AAATTOATET TTATTTTATATO TAAAQAQAET AEAQATTETT TETTOTOTOQ	12(
	STACATTAT CTAATCATAT TAAATATTAAA TATAAAAAT TOTTOAATA ATAATAATAA	18(
15	PTOTORTOLA COLLEGERA ARTESETO SARBARIO COLLEGERA ARTESETO ACCIAGIONO	24(
	TTTGTATAAC ACCTTCCCTTO OTOTCCCTTO OATTGGCATTAT	300
0.0	GACAAGCAAA TAAGAAAACC TIAGSTTTST TSTATTIGAA TYTYCGAAAAC AATAAAAAGGT	360
20	TINGACTCAA GATTIGCATI CAAGAAGAGG CAGAAATITT GICTVATCIT TITATCATTI	42.
	TGTGAACTTG TGTTTCTCTG TATGCTTAGA AAATTTTACA CACAAGGAAT GTTTGAAAAA	48(
25	CTGAGAATTT TAGAGTGCTT GGGTGGTTTT TATTTGGTCA GTGCTGAIGT GTTARGTGTT	540
	TAGGGAAATA ATGCTTCAGG ACCTTTTTGA CAACACAGYT TCATGAATGA CYGGGGGATA	600
20	TIWAKGITIGI GCIGAGAAAA CCGAGGAGI GGGCAGITIGG AATGGGGGAC CCITACCATI	66(
30	GGAAAACATG CATTCNGA.	67⊦
35	(2) INFORMATION FOR SEQ ID NO: 98: (1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1253 base pairs	
40	(B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linea:	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98	
٦.	ACCTCCCTCC CTCTCAGACT GGTCCGAATC CACGCCTAGC CCACCCACTG CCACTGGGGC	€(
	CATGEOCACO ACCACTG SG CACCGCCCC COTCOCTCC TOCCCACCCC	12(
50	CCTTECTCAG GCCCAGACCC AGCTEGGGCC CCACCGENAA GTTACCCCCA AGAGGCAAGI	18:
	NITGGCCTGA GACGCTCCTC ACTICTTAGA TCTTGGGGGC CTAAAGAGAC CCCCGTCCTG	24(
55	CCTCCTTTCT TTCTCTGTCT CTTCCTTTCCT TTTAGTCTTT TTCATCCTCT TCTCTTTCCA	30(
)'	CCAACCCTCC TGCATCCTTG CCTTGCAGCG TGACCGAGAT AGGTCATCAG CCCAGGGCTT	36(
	CAGTOTICCI TTATTTATAA TERRITUGUGG CTACCACCCA CCCTHCTUCA GTCTTGTGAA	42)
60	GAGTOTGAGA COTOCTTOTT COCCACTIOT CTCTTCCTTC ATTICTTUTT CTCTCCTTCT	481

	GOTOTOTI TINCTI ACAC TOTGACAGA ADIAATTATI ATTATTITTO TINTTOTITT	541
5	TYPTTYTTHER TYPTGERIES ARMCERATIC ETTIRARCRE ECCRETIENT ETTRUTY.	€((
2	ACAAAATATA MATATIYAAGA TGOTOOGTOO OCOTGTGAAG COCCCAGTGO CCCCGTGGG	660
	TONIAGICIES GERCOCATTO GESCCAAGOINE GATTOTOTOT ACCITACIACA CAGGCATGA	72
10	TOGGRATION OF TOTAL TOTAL SOCIETATION AND ACCORDANCE TO ACCUPANCE OF TOTAL TOT	761
	ACTOACTORS (POCASSESSIO) PROPERSION TOTOCOACCO CACCITICACA CACCITICACA	84(
15	ACTIVEACTIBE ATTOCASANT IGGA CANSTING CANAGEOTING CLIGOGGAAGG GCCCACTGC	90(
15	ARCHODERED SAIDAASEEF ITOODIDIAD DEDITIONIAD CODEASCOE TOTOODIDAA	96(
	ASSETTS STANAL ATTOCOURT AINTOCTICA STANGAS SESAS SESAS STANALEDEND	101(
20	ODACCOCET CTOCTAATCO TOCTCTCTCT TOATECTCASC TOCECTCEASCA ASAACAACAT	108(
	TIAGGERSA STATES STATE TATES AT A TOPE A TOTAGE A SA A SA SE A	114(
25	GATODADDAD TODODTDADD DOCOBADATD TMADENDODD AADDUUTAAD TOBATDTDTD	1200
۷.	TARAGGARIE GTIERCECTO ARARARARA ARARARARA ECTTGEGGGG GGG	1253
30	(2) INFORMATION FOR SEC ID NC: 99:	
30	(2) INFORMATION FOR SEQ ID NC: 99: (i) SEQUENCE CHARACTERISTICS: (A. LENGTH: 447 base pair: (B) TYPE: nucleic acic (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
35	(i) SEQUENCE CHARACTERISTICS: (A. LENGTH: 447 base pair: (B) TYPE: nucleic acic (C) STRANDEDNESS: double	
	(i) SEQUENCE CHARACTERISTICS: (A. LENGTH: 447 base pair: (B) TYPE: nucleic acic (C) STRANDEDNESS: double (D) TOPOLOGY: linear	€
35	(i) SEQUENCE CHARACTERISTICS: (A. LENGTH: 447 base pair: (B) TYPE: nucleic acic (C) STRANDEDNESS: double (D) TOPOLOGY: linea; (xi) SEQUENCE DESCRIPTION: SEQ IL NO: 99:	£:
35	(i) SEQUENCE CHARACTERISTICS: (A. LENGTH: 447 base pair: (B) TYPE: nucleic acic (C) STRANDEDNESS: double (D) TOPOLOGY: linea; (Xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99: CAAAGAATGA AATTUACCAC TOTCCTCTTC TTGGCAGGCTG TAGGAGGGGC CCIGGTCTAT	
35 40	(i) SEQUENCE CHARACTERISTICS: (A. LENGTH: 447 base pair: (B) TYPE: nucleic acic (C) STRANDEDNESS: double (D) TOPOLOGY: linea; (Xi) SEQUENCE DESCRIPTION: SEQ II: NO: 99: CAAAGAATGA AATTTACCAC TOTCCTCTTC TTGGCAGCTG TAGGAGGGGC CCIGGTCTAT GCTGAAGATG CCTCCTCTGA CTCGACGGGT GCTGATCCTN CCCAAGGAAGC TGGGACCTNC.	12(
35 40 45	(i) SEQUENCE CHARACTERISTICS: (A. LENGTH: 447 base pair: (B) TYPE: nucleic acic (C) STRANDEDNESS: double (D) TOPOLOGY: linear (Xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99: CARAGRATGA AATTVACCAC TOTCCTCTTC TTGGCAGCTG TAGGAGGGGC CCIGGTGTAT GCTGAAGRATG CCTCCTCTGA CTCGACGGGT GCTGATCCTN CCCAGGAAGC TGGGACCTNO: AAGCCTAATG AAGAGATCTC AGGTCCAGCA GAACCAGGTT CACCCCCAGA GACAACCAGT	12) 18)
35 40	(i) SEQUENCE CHARACTERISTICS: (A. LENGTH: 447 base pair: (B) TYPE: nucleic acic (C) STRANDEDNESS: double (D) TOPOLOGY: linear (Xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99: CARAGRATGA AATTTACCAC TOTCCTCTTC TTGGCAGCTG TAGGAGGGGC CCIGGTCTAT GCTGAAGRIG CCTCCTGGA CTGGACGGGT GCTGATCCTN CCCAGGRAGC TGGGACCTNT RAGGCTARTG RAGRIGATCTC AGGTCCAGCA GRACCAGGTT CACCCCCCAGA GACAACCAGT ACAGCCCAGG AGAYTTCCGC GGCAGCAGTT CAGGGGGACGG CCAAGGTCAC CTCAAGCAGT	12 (18) 24
35 40 45	(i) SEQUENCE CHARACTERISTICS: (A. LENGTH: 447 base pair: (B) TYPE: nucleic acic (C) STRANDEDNESS: double (D) TOPOLOGY: linear (Xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99: CARAGRATGA AATTVACCAC TOTCCTCTTC TTGGCAGCTG TAGGAGGGGC CCIGGTGTAD GCTGAAGRATGA CTCCTCTGA CTCGACGGGT GCTGATCCTN CCCAGGARAGC TGGGACCTNO: AAGCCTAATG AAGRGATCTC AGGTCCAGCA GAACCAGGTT CACCCCCCAGA GACAACCAGT ACAGCCCAGG AGAYTTCGGC GGCAGCAGTT CAGGGGACAG CCAAGGTCAC CTCAAGCAGC CAGGAACTAA ACCCCCTGAA AICCATAGTG GAGAAAAGTA TCTTACTAAC AGAACAAGC	12(18) 24 30

WO 98/39446 PCT/US98/04482

243

5	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 611 base pair: (E) TYPE: nucleic acic (C) STRANCEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:	
10	GETCTGGGGA GETGACATGI TGEGCTGTGG GATCCCAGCG CTGGGCCTGC TCCTGCTGCI	60
	GCAGGSWITC GCAGACGGAA ATGGAATCCA GGGATTCTTC TACCCATGGA GCTCTGAGGC	126
1.5	TGACATATYGG GACCGGGGAGA GCTCTGGGGGC CCGAGCGCC ATTCGATAGC CCCAACYTC	180
15	GCCTGCGTCI CCGGTGCTGC TACCGCAATG GGTCTGCTAC CACCAGCGTC CAGACGAAAA	240
	PARTOCTOR COTOCORDA COTOCACE TO TOUTOCOER TOTACACEAE CRADECTO	300
20	CTGCAGCATO TESTITEMINI COTGEGCCAA GEGECGGAC GIGCTGCATA TGCCCGGTTI	36C
	CCTGGCGGGI CCGTGTGACA TGTCCAAGTC CGTGTCGCTG CTCTCCAAGC ACCGAGGGA	420
0.5	TOTAGERADO TOADAAADOT DIDODOTRA OURTEDRADA DECADOTEON SOADAADAAD	480
25	ADAADDADD DDADADADAD DADTDDDAACD AEDDADDDAA DDECDADDDDA CEDADDDADDD	540
	AAAAAAAAA OAQQTQQOAT AQAAAAAAAAAAAAA TAQQAQQQ	60(
30	AAAAAAAA A	611
35 40	(2) INFORMATION FOR SEQ ID NO: 101: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 609 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 101:	
45	GCATTGCTAA AGCTGGCAGT TGAAACCAGT TGGACGCCC AGCTTGCGTC TCTTCTGCCT	60
	GAGTGGGCCT CTCAGGTCAC TCGTGCCCTG CTGGAGGACA GAGGGGCACC TCAGCCGCCC	120
÷0	CCAAGCCCAG AGCACAGCAA TAAGGTCG3C CT3CAGGAGC CGG3GTG3GG GTGGGGGTG.	180
50	GG3GRGCAGG ACCCTRARAT GCCACCAG3A CCT3ATGGGC CAG3AAGGGC GTGGACAT3C	240
	AGGCTGTTTT TACAGCTTTT TTTTTTTTGT TGTTTTTTTT TTAAAGAATA CAGAAGGAGG	300
55	CAAGCTTTTT TGCACTTTGT ATCCAGCTGC AAGCTCAGGG CAGAGTCAAG GGCCTGGGT	360
	GGAAAAACCT GACTCACAGG AATGCATAAT TGACCCTTGC AGCTACCCAA TAGCCCTTGC	42(

and the same

WO 98/39446 PCT/US98/04482

	CTORICA TOTOCOTORS COCCACACTO GREGOTOGAA GACTGOTTTO TABCACTACK	540
	GGTUACUGOC AGRICOTOST AGAAGATUC AGAAGATIAT TITACGTIGA COCCOTITITI	600
5	AATSTTCT1	€05
10	(C) ANTERNATION FOR SEC. IF NO. 166	
10	(2) INFORMATION FOR SEC ID NO: 102	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1770 base pairs	
15	(E) TYPE: nucleic acid (C) STHANDEDNESS: double	
	(D) TCPCLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEÇ ID NO: 102:	
20	ADDADIDERT BARTDAAART TAAARREDOUT ECROACODAR CHERROODIA AEROOVERDA	60
	MADAAINEWA EAGEBAARTO DENDITODEDA EGADAIRFER ADOSRAAAER REAGHAAERT	12(
25	GREETTING A GOGACOTTOS COSTRESSON GACCATOTT CTSCTSTOTS TOSTCACTAT	180
	CATCATCTGC TRUACCTGCT CCTGCTGCTG CCTTLACAAG ACGTGCCGCC GACCACGTCC	240
	GSTTSTCACC ACCACCACA CCACCACTEL GSTGCATGCC CCTTATCCTC AGCCTCCAAG	300
30	TSTSCOGOCO AGCTACCOTS GACCAAGCTA CCAGGGCTAC CACACCATGO CGCCTCAGC	360
	ACCORACION EMPONATTON ACCORACIONAL TODONACEMENT ACCORACIONAL TODONACEMENTA ACCORACIONAL TODONACEMENTA ACCORACIONAL TODONACEMENT ACCORACIONAL ACCORACIONAL TODONACEMENT ACCORAC	420
35	PADDROCCO TOODDODOO ACQUEENT OFFTDOORDA ROADDATEGO EGOASONEEQ	480
55	CCASCCTCCT TACAACCCGS SCTACATHRA THROCCCGAAG SORNCCTCTR ARCATTCCC	540
	GOCCTOINTS GOTGCCACTM GOTTATGED TONGTONES TWARTGGTON GOAGOOGG.	600
40	TTOOTTAGGO COCATGTGTG CURTGTGTGT CONGOCTGTA TATGTGGCTM COTOTGATG	660
	COUTOUTU STITCADADO ASSECTOSSEU CARACOSUTO CIAADAASSO SESDAAGADE	720
45	FEREFERRET SEERTHARRA ADTDAAASSE AASTTOTAAAA TOSTTOETAT TAAASTTOAD	780
40	CHICTERAGE TOTACAGE ACCTOTOTOGE REGREGAGE TOOCCEPTED ASIETOOCE	840
	TODASOTIAD CONASSEGA COMMENACIANOS CONSCIUSOS CASOTITUTS	900
50	COECTACORO ADTECTOR OPACOCATRO RACEROTURO ETOTECTATA DACETARIET	960
	CETTOTTTOTA GTIGGGGTA COTETTICAS AGCIAGGGAC ATSATGLAGS CGAAGITTO	1020
	GATETPECCA AGTTEGACT TEATCOTTELS GREATATETC CCATRECTOC CTGGAGCCTS	1080
55	TOATGOOTGT TERFEATCAG GOAGCOTOOT GATGOOAGAA CACCTOAGEC AGAGOOCTAC	1140
	TICASCIPETAC CIPETOTGECI GENETO CONTENDACA ETEROCOCE STOROCAGO	1200
60	GAGGGCACA TGCACACACA GOTTAGGTGC COCCAGGGAG CTCTGCTGCC CTCGGTGCC	1260

	CIGCCCITCC	CACAGGTGAG	CAGGGCTCCC	GTCCACCAGC	ACACTCAGTT	eremicaeme	1320
5	CYSLELLLIS	TETTATTTA	TAGAAAGGEA	TTPSSSSTTT	TTCTGTTTCA	AACATGATAG	1380
•	TTGATATGAG.	ACTGAAAACCC	CLECLECCLO	GAGGGAAATT	ADADAC/TOGO	TGGACAACCI	1440
	GGCAACTGTG.	ASTOCOTEST	TITITIGACATI	SEMACHICOSA	AATATGCAAC	APPTOOTSAA	1500
10	ACCTEACODO	CGSTGTTCTG	ASSERVEDED	CACCTEGGC	AATGGGCCAT	CTGGACCAAA	1560
	GELGGEELGT.	<u>өээөгсстээ</u>	AMBBCAGCMC	TESTCCAGAC	ATGAATACCT	CGTGTTCCT.	1620
15	TATOTOOOTO	TACTCTTTCA	CCAGAGTTGT	CTTAGCTCAA	ATCTGTTGTG	TTICTGAGTC	1680
1.	TAGGGTCTGI.	ACACTTGTTI	STAATAAATS	CAATCGTTTG	GAAAAAAAA	CAAAAAAAA	1740
	eeeeatsot	COATECCE	CAATSGCCTA.				1770
20							
	(2) INFORMA	TION FOR SE	eo id No. 10) ā .			
25							
2.2	(1)	(E) TYP	GTH: 1832 b E: nucleic	ase pairs acić			
0.0			ANDEDNESS: DLOGY: line				
30	(xi)	SEÇUENCE I	DESCRIPTION	: SEQ ID NO	: 103:		
	TASTESSTER	GTCATCTIGGA	GBAGATTTGC	TTTCTTTTTC	TCCAAAAGGG	GAGGAAATTG	€0
35	AAACTGCAGT	GGCCCACGAT	GGGAAGAGGG	GAAAGCCCAG	GGGTACAGGA	GBOOTOTGB'	120
	TGAAGGCAGA	GGCTAACATG	GGGTTTCGGAG	CGACCTTGGC	CGTTGGCTGA	CCATCTTTGT	180
40	GENGTOTGTC	GTCACTATCA	ТСАТСТЭСТТ	CACCTGCTCC	TGCTGCTGCC	TTTACAAGAC	240
40	GTGCCGCCGA	CCACGTCCGG	TYSTCACCAC	CACCACATCC	ACCACTGTGG	TGCATGCCCC	300
	TIATCCTCAG	OTTOCAAGTO	PACCOCOCCE	CTACCCTGGA	CCAAGCTACC	AGGGCTACCA	360
45	CACCATGCCG	CCTCAGCCAG	GGATGCCAGC	OATOOOAGGA	CCAATGCAGT	ACCCACCACC	420
	TTACCCAGCC	CAGCCCATGG	GCCACCGGC	CTACCACGAG	ACCCTGGCTG	GAGGAGCAGC	480
50	MATERCECER	COCGSCAGCC	AGCCTCCTTA	CAACCCGGCC	TACATGGATG	CCCGAAGCG3	540
50	COCTCTGAGC .	ATTCCCT/GGC	CTCTYTGSCT	GCCACTTGGT	TATETTETET	GTGTGCGTRA	600
	ASSTETEDTO	GGCGCGGTTC	CTTACGCCC	ATGTGTGCTG	ASSTETETET	CCCCCCCCC	660
55	CTTACGCCCC.	etteteteta	TOOMS ISTECT	TATATECTOCO	GTGGCTTCCT	CTGATGCTGA	720
	CAASTGGGGA .	PTTDDTAADA	CCEPTPAEACO	CTGGGACCAG	ACTITGTIGT	CTICCTCAC	78(
	TGAAATTATG	AAAATDOTTD	TOTCAAGOCA	AACTCAAAGA	ATGGGGTGGT	GGGGGCACC	84(

	CTOTOLAGOTO GOCCOTOLA LA GOLO KONNO DI TOTOCLAGOGO ACATOTONAS INTOTVOTOCA	900
	GOTTLACOCIA GUATGACCAA GUARRACUTU TUACACCAGRE GUGGUGARI TICURUGAGA	960
4	TGCAGATGTG TCCTGGTTTC GRUARYCETAG CCAGCTGCTA CTTGAGACAC TYGGCTCCTCC	102(
	COGSACTION RESTAUNCET TOWARABLOA GRANDAT GOAGROSAAG YITGASATO	1086
1C	AMADIRIDOD ARBIDIDIDID ITADDDTRIA EAUREWANTID DITARNITIDAD DITRAADDDD	1141
10	COTOTTOGGG ATCAGGEAGE OTCOMBATIGO CABACACOT CAGGACAGAGO COTACTCAGO	1200
	TGIACCTGTO TGCCTGAACT GICCCCTATC CCCGGATCTO CCCTAGAACC AGCTAGAAGGA	12€€
15	CCACATGOAC ACAMAGTOTA GOTGICOMER GGRAGOTOTE CTERROSTIFIC TIGGOCOTGO	1320
	CTTCCCACAG CTGACACACACACACACACACACACACACACACACACACA	138:
20	TITTCATTI ATTITACCOA AACATTITEC CESTITETES TIT SAAACAT GATASTEGAT	144
20	AAODDTO DAA DARDTAGADA OTODDTTAAA DERAGEMENT GEDTOGGGAA ADDOADA	1501
	ACCOCATECTO OTGATAGOTA TOOQACOAGA ECOCTTOGIC OCTGATAGOTO	1560
25	CTCCACGGTG TTCTGGCAGC AGGGACACCI GGGCCAATGG GCCATCIGGA CCAAAGGTGG	1620
	CONTONION TETROTICAL AADTAGAA COERTSTEER CERTACETON COERTSTEER	1680
30	TCTATTACIG TYTCACCAGA GCTGTCTIAG CICAAATCIG TYGIGTTTCT GAGTCTAGGG	1741
30	TCTGTACACT TGTTTATAAT AAANGCAATC GTTTNGGAAA AAAAANANAA AAAAAAAAAGG	1801
	GGSGGCGCTC TAAAAGGATN CCCCNAAGGG GG	1831
35		
	(2) INFORMATION FOR SEC IT NO: 104:	
40	(1) SEQUENCE CHAFACTEFISTICS:	
,,	(A) LENGTH: 2237 base pairs (B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double (D) TOPCLOSY: linear	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 104	
	AGTTCCCGGT ACTTTETTEC CARCETTGCC ATCHHARCE GHEATGACAI TACTCACTAI	€′
50	PARATTARA TOOCASOTA ORTRAGEST OSTAGALAST TEGETTAAAA OARTIAARAO	121
•	TAGTGTAAAG GATCTGAATG GCATASACTT AACTGCTGTG CAAGATACTC CIGTGGCTTC	180
	AAGAAAAGAA GATACATATG TTCATTTTAA TYYTYSACATT GAGCTCCAGA AGCATGTTGA	240
55	AAAATTAACC AAAGGTGCAG CIATCTTCTT TGAATTCAAA CACTACAAGC CIAAAAAAAG	300
	GTTTACCAGC ACCAACTGTT TTGCTTTCAT GGAGATGGAT GAAATTAAAC CTGGGCCAAT	360
60	TGTAATAGAA CIATACAAGA AACCCACTGA CTTTAAAAGA AAGAAATTGC AATTATTGAC	421

	CAAGAAACCA	CTITATICTTC	ATCTACATCA	AACTTTGCAC	AAGGAATGAT	CCTGACATGA	48
5	TGAACCTGGA	ACTIVITISTGA	ATTTIACCAC	TCAGTAGAAA	CCATCATAGC	TCTGTFTAG	54
-	ATATTCACCC	TTCAACAGGC	AGGAAGCAAG	SADDDATEDD	ACCACTAGGC	CGGATGGACT	60
	CAATNGCAAA	CACCATOTO	AGAATTICAGA	ADADE/ADOUTO	TOACACTGAC	GTATAGGACI	6+
10	CCTTGGGATA	CARRITTATI	GTAGATTTT\3	AAACATGTTT	TTACTTTTCT	ATTAATTGTV-	7: (
	CAATTAATAG	TOTATTATOT	AATTTACCAC	TACTOCTACC	CTGCTTCCTG	GAACAATACI	780
1 5	GTTGTG3G1A	GEATSTGOTO	ATCTTCAGAC	TIAATACACT	ARTAAGAATG	TGCTAGAGTT	64 1
	TACACATCTG	TTCACTTITG	CTCCAATATG	CTCTTTTGAC	TTAACGTCAA	GCTTTGGGT:	90
	GATGTGGGTA	GGSTAGTSTC	AAACTGCTTT	GAGAGGAATG	GBACCAGTTC	TGCTGCCTAA	<u>۶</u> (
20	GAAGGTCTGI	CTGGATGTTI	ATAGGCAGCA	CCTCTGAAGT	GGCCTAAATT	CACCCTGATC	101
	TGATAGTTTT	CCTGITIAGA	AAGTGTGCCT	TESCCAGATO	AGTATCCCAC	ATGGGAGTGT	108
25	TCCCTAGGTT	GTAGTTGTGA	TIGITTCCAG	ATGACCAGAT	TETTTTTCTG	AAAATGAGCA	1140
	PATTTTTAT	CATGTCGATI	AGCTGTTCTT	CTACATCACA	TTGTTACTCT	TTCTGATGAT	12(0
	GEDATOTTAG	TTAACATIGG	AACCATCTCA	CATTAATAAA	AAAGTTTTAG	ATGGGTTTA(12t
30	AATGTCTTCI	AAACAATGTA	ATCTAAAAAT	AATIGAGTCA	GATGCTAACG	AGATACTGCA	131
	GGCATAACTG	CTSTTTTCT	GACAACTGAT	TGTGAAACCT	TAAAACCTGC	ATACCTCTT	13F
35	TTACAGTGAG	GACTATGCAA	AATCTGGAAA	GATATTCTAT	ATATTTTTTT	TAGGTAGATA	144
	GGATCGCCAT	TTATITCCTA	TTTAGATATA	CTGACATTCA	TOCATATGAA	AATATGCAG:	1500
	TCATTAGCTT	ACTATAATTT	ACTITIGACT	TAATG3GGCA	TAAATAAAAC	TTTCATAGTA	1560
10	CACATGAGGT	GGATATTPGA	TACACAGAAC	ATTTGCGGTG	GGCTTTCTGT	GGGTTAGATV:	1610
	TAAAGCCCAC	TAATTTTAAT	ATTCACTATT	TIAAATGAGC	AATGCATGAG	GGGAATGCAC	1 <i>€</i> :E-
15	TGTCAGTACC	TGGCCTATTT	TTAAACTAGT	GTAATCACCC	TAGTCATACC	ATTCAGTAT	174
	TTTGCTTTTT	AAAATAAGTA	ACCACAATTA	AFTTGTTGTA	GCCTTGCAC	TTCAAGAGAT	180
	CTAGTCTTTA	CTTTCAGTTG	TCTGTTAGGT	CCATTCTGTI	TACTAGACGG	ATGTTAATAA	1860
50	AAACTATGCG	AGCCTGAATG	AATTCTCAGC	CAAATTTAGT	CTTGTCTCTC	ATCTTGATTG	1921
	GATTAATTCC	AAATTCTAAA	ATGATTCAGT	CCACAATAGC	TCTAGGGGAT	GAAGAATTTG	1980
55	CCTTACTTTG	CCCASTTCCT	AAGACTGTGA	GTTGTCAAAT	CCCTAGACTG	TAAGCTCTTC	2040
-	AAGGAGCAAG	AGGOGCATTT	TOTOCGTGTC	TTTTAATOIA	TOTAAGGTGT	TTGGCAGCAC	2100
	TCTGTACCCT	GIGGAGIACI	CAGTACCTTT	TETTTGATGT	TGCTGACAAG	ACCTGAAAA	216
50	AAATCCCTTA	AAAAAAAA	CCATTAAAGT	GTAGCAAAAC	AAAAAAA	AAAANAAAAA	2220

PCT/US98/04482 WO 98/39446

248

2237 ACTOGAGACG GGCCCGC

5

(2) INFORMATION FOR SEQ ID NO 105:

(i) SEQUENCE CHARACTERISTICS. 10 (A) LENGTH: 1822 base pairs (E) TYPE: nucleic acid

(C) STRANDEDNESS: double (D) TOPOLOGY linear 15 (xi) SEPWENCE DESCRIPTION: SEQ ID NO: 105: GGTCGACCCA CGCGTCCGGA ATTITCCGTAG CAATAAGTTT GTGCATGTAT AGTAATTIG 60 ATTAGCAAGG TICTAACCTY TGCCTCTTGG GITCAAGTGA TTCTCGTGCC CCAGCCTCCC 120 20 GACTAGOTYS GACTACAGSC ACCTCCCACO ACCTCCAGOT AATTTTTATA TTYTTAGTAG 180 AGACEGETT TTOUTETETT GECCAGGCTG GICTCAAACT CCTGACCTCA AGTAATCCAC 240 300 25 CTGGCCTGCT CTTTCATGT CTTAACATGG CATGTCTTTT AGTTTCATTA TTTTCCTACT CCTTSTATGT CAAGAAATTA CATTTTGCAT GTCTTATGGA GATGCTGTTA ATTGCTTCAG 360 TGAGTGCTTT TCTAATCTGC AGACCATTTA CATTTCCTGT TTGCAGCATG CTGTGTGCAA 420 30 ACACTCAGIA ATTIGGAGIA TICAATTATI TOTTAGGGCT CTICCTATIT CCAAATGIG! 480 TGAATTGTCT ATVGATGGGA TTTTCAGATC TVTTCATGAG AACTGGAAAT GTAGCTGGGT 540 GGCACCTACC TAGGITGCTA CGTAGTGAGT AGACTTTCTC TTGGGTATAG TAAGCCTCAG 600 35 ACAGCTITCA CITTIATUTA CITTACITGI GUAAATAAAA CAGTCATITT GITCIGAAAG 660 AATAAGATAG CTTTCTGTAG AGAAGGAATT CCTACCTCTA AAAGCTGCCT TGAGAACTCA 40 GAACTGGCAG TTTTCTGAGG TGATTTTTAA ATTTCAGTAT TAGGGAGAGT CCAGCATTTG 780 CTGACACAGA TTCTACATAA CTAATGTATG ATAGCAAATG CAAAACTATT ATAATGTGG: 840 900 45 GTATCTTGTG CATACACAGG TTAGAACAAG TAGACTCTGG CAGCAGATCT CCAGAGACCC 960 AAGTTTAGGT TOTCATAGTG TATTTGAAGT AGTTATACTC CTGGCTTAAG TAGTTTAGTG CCTGGGAGAA TOCATTACTG AAAAGCATTI AACTTAAAAA AAAAAAAAA AAAAAAAAAA 1020 50 AAACCTCGTG COGAATTCGG CACGAGCTAA CCCAGAAACA TCCAATTCTC AAACTGAAGC 1080 TOGOACTOTO GOOTCOAGOA IGAAAGTOTO TGOOGCOOTT OTGIGOCTGO TGOTCATAGO 1140 AGCCACCTTC ATTCCCCAAG GGCTCGCTCA GCCAGATGCA ATCAATGCCC CAGTCACCTG 1200 55 CTGYTATAAC TYCACCAATA GGAAGATCTC ACTGCAGAGG CTCGCGAGCT ATAGAAGAAT 1260 CACCAGCAGC AAGTGTCCCA AAGAAGCTCT GATCTTCAAG ACCATTGTGG CCAAGGAGAT 1320

	CTGTGCTGAC	CCCAAGCAGA	AGTGGGTTCA	GGATTCCATG	GACCACCTGG	ACAAGCAAAC	1380
	CCAMACTCCG	AAGACTTGAA	CACTCACTCC	ACAACCCAAG	AATCTGCAGC	TAACTTATTI	1440
5	TCCCCTAGCT	TTCCCCAGAC	ACCCTGTTTT	TATTATTTTA	AATGAATTTT	GTTTGTTGAT	1500
	GTGAAACATT	ATGCCTTAAG	TAATTSTTAAT	AATTTATTOT	GTTATTGATG	TTTTAAGTTI	1560
10	ATCTITICATG	GTACTAGTGT	TTTTT AGATA	CAGAGACTTG	GGGAAATTGC	TTTTCCTCTT	1620
10	GAACCACAGT	TCTACCCCTG	GEATETTTTG	AGGGTCTTTG	CAAGAATCAT	TAATACAAAG	1680
	AATTTTTTTT	AACATTCCAA	TGCATTGCTA	TTATTATAAA	GTGGAAATGA	ATATTTTGTA	1740
15	CACATTAICA	CAAATAAATA	TATITITGIA	САААААААА	AAAAAAAA	AAAAAAAA	1800
	AAGSGGCCGC	TCGAATTAAG	CC				1822

20

25

The state of the second

(2) INFORMATION FOR SEQ ID NO: 10ϵ :

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1712 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: dcuble

(D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:

	CGIGCCCCAG	CCTCCCGAGT	AGCTGGRACT	ACAGGCACGT	SCCACCACGC	CCAGCTAATT	60
35	TTWATATTT	WAGTAGAGAC	GEGTTTTSC	TSTKTTGGCC	AGGCTGGTCT	CAAACTCCTG	120
	ACCTCAAGTA	ATCCACCTGG	CCTGCTCTTT	TCATGTCTTA	ACATGGCATG	TCTTTTAGTT	180
	TCATTATTTT	CCTACTCCTT	GTATGTCAAG	AAATTACATT	TTGCATGTCT	TATGGAGATG	240
40	CTGTTAATTG	CTTCAGTGAG	TGCTTTTCTA	ATCTGCAGAC	CATTTACATT	TCCTGTTTGC	300
	AGCATGCTGT	GTGCAAACAC	TCAGIAATTI	GGAGTATTCA	ATTATTTGTT	AGGGCTCTTC	360
45	CTATTTCCAA	ATGTGCTGAA	TTGTCTATTG	ATGGGATTTT	CAGATCTTTT	CATGAGAACT	420
	GGAAATGTAG	CTGGGTGGCA	CCTACCTAGG	TTGCTACGTA	GTGAGTAGAC	TTTCTCTTGG	480
	GTATAGTAAG	CCTCAGACAG	CTTTCACTTT	TATCTACTTT	ACTTGTGGAA	ATAAAACAGT	540
50	CATTTTGTTC	TGAAAGAATA	AGATAGCTTT	CTGTAGAGAA	GGAATTCCTA	CCTCTAAAAG	600
	CTGCCTTGAG	AACTCAGAAC	TYGGCAGTTTT	CTGAGGTGAT	TTTAAATTT	CAGTATTAGG	660
55	GAGAGTCCAG	CATTTGCTGA	CACAGATTCT	ACATAACTAA	TGTATGATAG	CAAATGCAAA	720
	ACTATTATAA	TGTGGTGTAT	CTTGCGCATA	CACAGGTTAG	AACAAGTAGA	CTCTGGCAGT	780
	AGATCTCCAG	AGACCCAAGT	TIAGGTTCTC	ATAGTGTATT	TGAAGTAGTT	ATACTCCTG3	840
60	CTTAAGTAGT	TTAGTGCCTG	GGAGAATCCA	TTACTGAAAA	GCATTTAACT	AAAAAAAT	900

	CTCCTGCCGA ATCGGCCGA ATTCGGCAGAAAAAAAAA TCCAATTCTC	961
_	DETERMINE TERRESPORTE STEERANDE AUGUSTONE SECONDE DESARANTE	102:
5	TOCTCATAGO AGCCACCITO ATTOCCCAGA GACCCCCAGCAGACACA ATCAATACC	108/
	CAGTCACCTG CTGYTATAAC TTCACCAATA GGAAGATCTC AGTGCAGGG CTCGCGAGCT	1140
0	ATAGAAGAAT CACCAGCAGC AASTSTCCCA AAGAAGCTGT GATSTTCAAG ACCATTGTGG	1200
	CCAAGGAGAT CTGTGCTGAC CCCAAGCAGA AGTGGGTTCA GGATTCCATG GACCACCTGG	1260
	ACAAGCAAAC CCAAACTCG AAGATTAGA CACTCACTC ACAACCAAG AATCTGCAGC	1320
15	TETTAADUAA TATTATUTA TETTOTOOGA OAEACCCCTT IODATOCCC TTTATAADUAA	1380
	CTTTGTTGAT GTGAAACATT ATGCCTTAAG TAATGTTAAT TCTTATTTAA GTTATTGATG	1440
20	TTTTAAGTT ATCTTTCATG GIACIAGTGI TTTTTAGATA CAGAGACTTG GGGAAATTGC	1500
	TTTTCCTCTT GAACCACAGT TCTACCCCTG GGATGTTTTG AGGGTCTTTG CAAGAATCAT	1560
	TAATACAAAG AATTTTTTT AACATTCCAA TGCATTGCTA AAATATTATT GTGGAAATGA	1620
25	ATATTTTGTA ACTATTACAC CAAATAAATA TATTTTTTGTA CAAAAAAAA AAAAAAAAAA	1680
	AAAAAAAA AAGSGGCCGC TCGAATTAAG CC	1712
30		
	(2) INFORMATION FOR SEQ ID NO: 107:	
35	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 1969 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
40	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 107:	
	CCCCTCCTTC CCCTYGCCAC CTACTGAACC CTCCTCCGAG GTGCCCGAGC AGCCGTCTGC	60
45	CCAGCCACTC CCTGGGAGTC CCCCCAGAAG AGCCTATTAC ATCTACTCCG GGGCCGAGAA	120
	GATCCCCCTG GTGTTGAGCC GGCCCCTCTC CTCCAACGTG GCCACTCTTC AGCATCTCTG	180
50	TCGGAAGACC GTCAACGGCC ACCTGGACTC CTATGAGAAA GTCACCCAGC TGCCGGGGCC	240
50	CATTCGGRAG TTCCTGGACC AGTACGATGC CCCGMTTTAA GGGGTAAAGG GCGCAAAGGG	300
	CATGGGTCGG GAGAGGGGAC GCAGGCCCCT CTCCTCCGTG GCACATGGCA CAAGCACAAG	360
55	AAGCCAACCA GGAGAGAGTC CTGTAGCTCT GGGGGGAAAG AGGGCGGACA GGCCCCTCCC	420

GTACCACCTG AGTCTCCAGC TTCTCCGGAG ACCCAGCTGT CCTGGTGGGA CGATAGCAAC

	CACAAGTGGA	TICTCCTTCA	ATTCCTCAGC	TTCCCCTCTG	CCTCCAAACA	GGGGACACTT	600
	CGGGAATGCT	GAANTAATGA	GAACTGCCAG	GGAATCTTCA	AACTTTCCAA	CGGAACTTGT	660
5	TIGCTCTTIG	ATTTGGTTTA	AACCIGAGCT	GGTTGTGGAG	CCTG3GAAAG	GTGGAAGAGA	72(
	GAGAGGTCCT	GAGGGCCCA	GGGSTGCG3G	CTGGCGAAGG	AAATGGTCAC	ACCCCCCCCC	781
10	CACCCCAGGC	GAGGATCCTG	CTGACATGCT	CCTCTCCCTG	GCTCCGGGGA	GAAGGGCTTG	840
10	GGGTGACCTG	AAGGGAACCA	TCCTGGTGCC	CCACATCCTC	TCCTCCGGGN	ACAGTCACCG	900
	AAAACACAGG	TTCCAAAGTC	TACCTGGTGC	CTGAGAGCCC	AGGGCCCTTC	CTCCGTTTTA	96(
15	AGGGGGAAGC	AACATTTGGA	GGGGACGGAT	GGGCTGGTCA	GCTGGTCTCC	TTTTCCTACI	102(
	CATACTATAC	CTTCCTGTAC	CTGGGTGGAT	GGAGCGGGAG	GATGGAGGAG	ACGGGACATC	108(
20	TTTCACCTCA	GGCTCCTGGT	AGAGAAGACA	GGGGATTCTA	CTCTGTGCCT	CCTGACTATG	1140
20	TCTGGCTAAG	AGATTCGCCI	TAAATGCTCC	CTGTCCCATG	GAGAGGGACC	CAGCATAGGA	1200
	AAGCCACATA	CTCAGCCTGG	ATGGGTGGAG	AGGCTGAGGG	ACTCACTGGA	GGGCACCAAG	126(
25	CCAGCCCACA	GCCAGGGAAG	TGGGGAGGG	GGGCGGAAAC	CCATGCCTCC	CAGCTGAGCA	132(
	CTGGGAATGT	CAGCCCAGTA	AGTATTGGCC	AGTCAGGCGC	CTCGTGGTCA	GAGCAGAGCC	1380
30	ACCAGGTCCC	ACTGCCCCGA	GCCCTGCACA	GCCCTCCCTC	CIGCCIGGGT	CGGGGAGGCT	144(
50	GGAGGTCATT	GGAGAGGCTG	GACTGCTGCC	ACCCCGGGTG	CTCCCGCTCT	GCCATAGCAC	1500
	TGATCAGTGA	CAATTTACAG	GAATGTAGCA	GCGATGGAAT	TACCTGGAAC	ATTTTTTGTT	15€€
35	TTTGTTTTTG	TTTTTGIITT	TCTGGGGGGG	GGCAACTAAA	CAAACACAAA	GTATTCTGTG	1620
	TCAGGTATTG	GGCTGGACAG	GGCAGTTGTG	IGTTGGGGTG	GTTTTTTCT	CTATTTTTT	168(
40	GTTTGTTTCT	TGTTTTTAA	TAATGTTTAC	AATCTGCCTC	AATCACTCTG	TCTTTTATAA	1740
40	AGATTCCACC	TCCAGTCCTC	TCTCCTCCCC	CCTACTCAGG	CCCTTGAGGC	TATTAGGAGA	1800
	TGCTTGAAGA	ACTCAACAAA	ATCCCAATCC	AAGTCAAACT	TTGCACATAT	TATTTATATT	1860
45	ATTCAGAAAA	GAAACATTTC	AGTAATTTAT	' AATAAAGAGC	ACTATTTTT	AATGAAAAA	192(
	AAAAAAAA	AAAAAAAA	CGACGCTGGT	GACCGGAATY	CGACGTACG		19€9

55

to to the constant A.

(2) INFORMATION FOR SEQ ID NO: 108:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1734 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:

	CGGGTCCCAA	GOCTGTGOCT	GAGCCTGAGC	CIGAGCCTGA	GOTOBAGOOG	GGAGCCGGTY	€(
5	GEGGGGGETE	CBBBCTSTBG	GACCGCTGGG	ODDO DAGOGA	TGGGACCCT	GTGGGGAGGT	120
J	CTTCTTCGGC	TTGGTGGTT	GCT/CAGCCT/G	TOGTSCOTGS	CGCITTCCGT	GCTGCTGCTV	180
	GUGCATYGTING	AGACGCCGCC	AAGAATTTOG	AGGATGTCAG	ATGTAAATGT	ATCTGCCCTC	14(
10	CCTATAAAGA	AAATTCT9GG	CATATTTATA	ATAAGAACAT	ATCTCAGAAA	GATTGTGATT	300
	GCCTTCATGT	TGTGGAGCCC	CETTOTOCTA	ADTOCEDEED	TOTAGAAGCA	TACTGTCTAC	360
15	GCTGTGAATG	САААТАТЗАА	GAAAGAAGCT	CTGTCACAAT	CAAGGTTACC	PTTAATATTA	420
15	ATCTCTCCAT	TTTGGGGCTT	CTACTTCTGT	ACATGGTATA	TOTTACTOTG	GTTGAGCCCA	480
	TACTGAAGAG	GOGOGTOTTT	GGACATGCAC	AGTTGATACA	GAGTGATGAT	GATATTGGGG	540
20	ATCACCAGCC	TITTGCAAAT	GCACACGATG	TGCTAGCCCG	CTCCCGCAGT	CGAGCCAACG	600
	TGCTGAACAA	GSTAGAATAT	GCACAGCAGC	GCTGGAAGCT	TCAAGTCCAA	GAGCAGCGAA	660
25	AGTCTGTCTT	TGACCGGCAT	GTTGTCCTCA	GCTAATTGGG	GAATTGAATT	CAAGGTGACT	720
23	AGAAAGAAAC	AGGCAGACAA	CTGGGAAAGA	ACTGACTGGG	NTTTTGCTGG	GTTTCATTTT	760
	AATACCTTGT	TGATTTCACC	AACTGTTGCT	GGAAGATTCA	AAACTGGAAG	CAAAAACTTG	840
30	CTTGATTTTT	TTTTCTTGTT	AACGTAATAA	TAGAGACATT	TITAAAAGCA	CACAGCTCAA	900
	AGTCAGCCAA	TAAGTCTTTT	CCTATTTGTG	ACTITITACTA	AATAAAATA	ATCTGCCTGT	960
35	AAATTATCTT	GAAGTCCTTT	ACCTGGAACA	AGCACTCTCT	TTTTCACCAC	ATAGTTTTAA	1020
<i>JJ</i>	CTTGACTTTC	AAGATAATTT	TCAGGGTTTT	TGTTGTTGTT	GTTTTTTGTT	TGTTTGTTTT	1080
	GGTGGGAGAG	GGGAGGGATG	CCTGGGAAGT	GGTTAACAAC	TTTTTTCAAG	TCACTTTACT	1140
40	AAACAAACTT	TTGTAAATAG	ACCTTACCTT	CTATTTTCGA	GTTTCATTTA	TATTTTGCAG	1200
	TGTAGCCAGC	CTCATCAAAG	AGCTGACTTA	CTCATTTGAC	TTTTGCACTG	ACTGTATTAT	1260
45	CTGGGTATCT	GCTGTGTCTG	CACTTCATGG	TAAACGGGAT	CTAAAATGCC	TGGTGGCTTT	1320
73	TCACAAAAAG	CAGATTTTCT	TCATGTACTG	TGATGTCTGA	TGCAATGCAT	CCTAGAACAA	1380
	ACTGGCCATT	TGCTAGTITA	CTCTAAAGAC	TAAACATAGT	CTTGGTGTGT	GTGGTCTTAT	1440
50	TCATCTTCTA	GTACCTTTAA	GGACAAATCC	TAAGGACTTG	GACACTTGCA	ATAAAGAAAT	1500
	TTTATTTTAA	ACCCAAGCCT	CCCTGGATTG	ATATATATA	CACATTTGTC	AGCATTTCCG	1560
55	GTCGTGGTGA	GAGGCAGCTG	TTTGAGCTCC	AATGTGTGCA	GCTTTGAACT	AGGGCTGGGG	1620
55	TIGTGGGTGC	CTCTTCTGAA	AGGTCTAACC	ATTATTGGAT	AACTGCTTT	TTTCTTCCTC	1680
	TTTGGAATGT	AACAATAAAA	ATAATTTTTG	AAACATCAAA	AAAAAAA	AAAA	1734
60							

	(2) INFORMA	TION FOR SE	Ç ID NO: 10	·¢ :			
5	(<u>i</u>)	(E) TYPE (C) STRA	MARACTERISTI STH: 2003 ba E: nucleic a ANDEDNESS: c DLOGY: linea	ase pairs acid double			
10	(xi)	SEQUENCE I	DESCRIPTION:	: SEÇ ID NC:	109:		
	CGCAGGGGGC	9000690600	GGGACTCGCA	TTCCCCCGGTT	CCCCCTCCAC	CCCACGCGGC	6(
15	CTGGACCATG	GACGCCAGAT	GFTGGGCAGT	GGTGGTGCTG	GITGOGTTCC	CCTCCCTAGG	12(
	GGCAGGTGGG	GAGACTCCCG	AAGCCCCTCC	GGAGTCATGG	ACCCAGCTAT	GSTTCTTCCG	18(
20	ATTIGTGGTG	AATGCTGCTG	GTTATGCCAG	NITTATGGTA	CCTGGCTACC	TCCTGGTGCA	240
20	GIACTICAGG	CGGAAGAACT	ACCTGGAGAC	CGGTAGGGGC	ChClesinic	CCCTGGTGAA	300
	AGCTTGTGTG	TTIGGCAATG	DDDAADDDDAA	CTCTGATGAG	GTTCCCCTGG	CGCCCCGAAC	361
25	AGAGGCGGCA	COCACCAGA	CGATGTGGCA	GGCCCTGAAG	CTGCTCTTCT	GTGCCACAGG	420
	GCTCCAGGTG	TOTTATOTGA	CTTGGGGTGT	GCTGCAGGAA	AGAGTGATGA	CCCGCAGCTA	481
30	TGGGGCCACA	OCCACATO	CGGGTGAGCG	CTTTACGGAC	TYGGAGTTCC	TGGTGCTAAT	54
50	GAACCGAGTG	CTGGCACTGA	TTGTGGCTGG	CCTCTCCTGT	GTTCTCTGCA	AGCAGCCCCG	600
	GCATGGGGCA	CCCATGTACC	GGTACTCCTT	TGCCAGCCTG	TOCAATGTGC	TTAGCAGCTG	660
35	GTGCCAATAC	GAAGCTCTTA	AGTTCGTCAG	CTTCCCCACC	CAGGTGCTGG	CCAAGGCCTC	720
	TAAGGTGATC	CCTGT CATGC	TGATGGGAAA	GCTTGTGTCT	CGGCANTA	ACGAACACTG	780
40	GGAGTACCTG	ACAGCCACC	TCATCTCCAT	TGGGGTCAGC	ATGTTTCTGC	TATCCAGCGG	840
40	ACCAGAGCCC	CGCAGCTCCC	CAGCCACCAC	ACTCTCAGGC	CTCATCTTAC	TGGCAGGTTA	90(
	TATTGCTTTT	GACAGCTTCA	CCTCAAACTG	GCAGGATGCC	TETTTGCCTA	TAAGATGTCA.	960
45	TCGGTGCAGA	TGATGTTTGG	GGTCAATTTC	TTCTCCTGCC	TETTCACAGT	GGGSTCACTG	1020
	CTAGNAACAG	GGGGGMCCTA	CTGGAGGGAA	CCCGCTTCAT	GGGGCGACAC	AGTGAGTTTG	1080
50	CTGCCCATGC	CCTGTTACTC	TCCATCTGCT	CCGCATGTGG	CCAGCTCTTC	ATCTTTTACA	1140
30	CCATTGGGCA	GTTTGGGGCT	GCCGTCTTCA	CCATCATCAT	GACCCTCCGC	CAGGCCTTTG	1200
	CCATCCTTCT	TTCCTGCCTT	CTCTATGGCC	ACACTGTCAC	TGTGGTGGGA	GGGCTGGGGG	1260
55	TGGCTGTGGT	CTTTGCTGCC	CTCCTGCTCA	GAGTCTACGC	GCGGGGCCGT	CTAAAGCAAC	1320
	GGGGAAAGAA	GGCTGTGCCT	GTTGAGTCTC	CTGTGCAGAA	GGTTIGAGGG	TGGAAAGGGC	1380

CTGAGGGGTG AAGTGAAATA GGACCCTCCC ACCATCCCCT TCTGCTGTAA CCTCTGAGGG 1440

TANADA PARAMANAN

	AGCTGGCTGA AAGGGCAAAA TGCNGGTGTT TTCTCAGTAT CACAGACCAG CTCTGCAGCA	150C
	GGGGATTYBYG GAYCCCAGGA GGCAYCCTTC CCTTTTBCCT TAAGTCACCC ATCTTCCAGT	1560
5	AAGCAGTTTA TTTTGAGCCC CGGGGGTAGA CAGTCCTCAG TGAGGGGTTT TGGGGAGTTT	1€20
	GGGGTCAAGA GAGCATAGGT AGGTTCCACA GTYACTCTTC CCACAAGTTC CCTTAAGTCT	168C
10	TGCCCTASCT CTSCTCTGCC ACCTTCCAGA CTCACTCCCC TCTGCAATA CCTGCATTTC	1740
10	TTACCCTGGT GAGAAAAGCA CAAGCGGTGT AGGCTCCAAT GCTGCTTTCC CAGGAGGGTG	1800
	AAGATGGTGC TGTGCTGAGG AAAGGGGATG CAGAGCCCTG CCCAGCACCA CCACCTCCTA	1860
15	TGCTCCTGGA TCCCTAGGCT CTGTTCCATG AGCCTGTTGC AGGTTTTGGT ACTTTAGAAA	1920
	TGTAACTTYT TGCTCTTATA ATTTTATTT ATTAAATTAA ATTACTGCAA AAAAAAAAA	1980
20	AAAAAATCG GGGGGGGCC CGN	2003
	(2) INFCRMATION FOR SEQ ID NO: 110.	
25	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 1320 base pairs (B) TYPE: nucleic acid	
30	(C) STRANDEDNESS: doubl∈ (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 110:	
25	GCTGAGCTGC CTTGAGGTGC AGTGTTGGGG ATCCAGAGCC ATGTCGGACC TGCTACTACT	60
35	GCTGAGCTGC CTTGAGGTGC AGTGTTGGGG ATCCAGAGCC ATGTCGGACC TGCTACTACT GGGCCTGATT GGGGGCCTGA CTCTCTTACT GCTGCTGACG CTGCTGGCCT TTGCCGGGTA	60 120
35		
35 40	GGGCCTGATT GGGGGCCTGA CTCTCTTACT GCTGCTGACG CTGCTGGCCT TTGCCGGGTA	120
	GGGCCTGATT GGGGGCCTGA CTCTCTTACT GCTGCTGACG CTGCTGGCCT TTGCCGGGTA CTCAGGGCTA CTGGCTGGGG TGGAAGTGAG TGCTGGGTCA CCCCCCATCC GCAACGTCAC	120 180
40	GGGCCTGATT GGGGGCCTGA CTCTCTTACT GCTGCTGACG CTGCTGGCCT TTGCCGGGTA CTCAGGGCTA CTGGCTGGGG TGGAAGTGAG TGCTGGGTCA CCCCCCATCC GCAACGTCAC TGTGGCCTAC AAGTTCCACA TGGGGCTCTA TGGTGAGACT GGGCGGCTTT TCACTGAGAG	120 180 240
	GGGCCTGATT GGGGGCCTGA CTCTCTTACT GCTGCTGACG CTGCTGGCCT TTGCCGGGTA CTCAGGCCTA CTGGCTGGGG TGGAAGTGAG TGCTGGGTCA CCCCCCATCC GCAACGTCAC TGTGGCCTAC AAGTTCCACA TGGGGCTCTA TGGTGAGACT GGGCGGCTTT TCACTGAGAG CTGCAGCATC TCTCCCAAGC TCCGCTCCAT CGCTGTCTAC TATGACAACC CCCACATGGT	120 180 240 300 360
40	GGGCCTGATT GGGGGCCTGA CTCTCTTACT GCTGCTGACG CTGCTGGCCT TTGCCGGGTA CTCAGGCCTA CTGGCTGGGG TGGAAGTGAG TGCTGGGTCA CCCCCCATCC GCAACGTCAC TGTGGCCTAC AAGTTCCACA TGGGGCTCTA TGGTGAGACT GGGCGGCTTT TCACTGAGAG CTGCAGCATC TCTCCCAAGC TCCGCTCCAT CGCTGTCTAC TATGACAACC CCCACATGGT GCCCCCTGAT AAGTGCCGAT GTGCCGTGGG CAGCATCCTG AGTGAAGGTG AGGAATCGCC	120 180 240 300 360
40	GGGCCTGATT GGGGGCCTGA CTCTCTTACT GCTGCTGACG CTGCTGGCCT TTGCCGGGTA CTCAGGCCTA CTGGCTGGGG TGGAAGTGAG TGCTGGGTCA CCCCCCATCC GCAACGTCAC TGTGGCCTAC AAGTTCCACA TGGGGCTCTA TGGTGAGACT GGGCGGCTTT TCACTGAGAG CTGCAGCATC TCTCCCAAGC TCCGCTCCAT CGCTGTCTAC TATGACAACC CCCACATGGT GCCCCCTGAT AAGTGCCGAT GTGCCGTGGG CAGCATCCTG AGTGAAGGTG AGGAATCGCC CTCCCCTGAG CTCATCGACC TCTACCAGAA ATTTGGCTTC AAGGTGTTCT CCTTCCCGGC	120 180 240 300 360 420
40 45	GGGCCTGATT GGGGGCCTGA CTCTCTTACT GCTGCTGACG CTGCTGGCCT TTGCCGGGTA CTCAGGGCTA CTGGCTGGGG TGGAAGTGAG TGCTGGGTCA CCCCCCATCC GCAACGTCAC TGTGGCCTAC AAGTTCCACA TGGGGCTCTA TGGTGAGACT GGGCGGCTTT TCACTGAGAG CTGCAGCATC TCTCCCAAGC TCCGCTCCAT CGCTGTCTAC TATGACAACC CCCACATGGT GCCCCCTGAT AAGTGCCGAT GTGCCGTGGG CAGCATCCTG AGTGAAGGTG AGGAATCGCC CTCCCCTGAG CTCATCGACC TCTACCAGAA ATTTGGCTTC AAGGTGTTCT CCTTCCCGGC ACCCAGGCAT GTGGTGACAG CCACCTTTCCC CTACACCACC ATTCTGTCCA TCTGGCTGGC	120 180 240 300 360 420
440 445 50	GGGCCTGATT GGGGGCCTGA CTCTCTTACT GCTGCTGACG CTGCTGGCCT TTGCCGGGTA CTCAGGCCTA CTGGCTGGGG TGGAAGTGAG TGCTGGGTCA CCCCCCATCC GCAACGTCAC TGTGGCCTAC AAGTTCCACA TGGGGCTCTA TGGTGAGACT GGGCGGCTTT TCACTGAGAG CTGCAGCATC TCTCCCAAGC TCCGCTCCAT CGCTGTCTAC TATGACAACC CCCACATGGT GCCCCCTGAT AAGTGCCGAT GTGCCGTGGG CAGCATCCTG AGTGAAGGTG AGGAATCGGC CTCCCCTGAG CTCATCGACC TCTACCAGAA ATTTGGCTTC AAGGTGTTCT CCTTCCCGGGC ACCCAGGCAT GTGGTGACAG CCACCTTTCCC CTACACCACC ATTCTGTCCA TCTGGCTGGC TACCCGCTGT GTCCATCCTG CCTTGGACAC CTACATCAAG GAGCGGAAGC TGTGTGCCTA	120 180 240 300 360 420 480 540 600
40 45	GGGCCTGATT GGGGGCCTGA CTCTCTTACT GCTGCTGACG CTGCTGGCCT TTGCCGGGTA CTCAGGCCTA CTGGCTGGGG TGGAAGTGAG TGCTGGGTCA CCCCCCATCC GCAACGTCAC TGTGGCCTAC AAGTTCCACA TGGGGCTCTA TGGTGAGACT GGGCGGCTTT TCACTGAGAG CTGCAGCATC TCTCCCAAGC TCCGCTCCAT CGCTGTCTAC TATGACAACC CCCACATGGT GCCCCCTGAT AAGTGCCGAT GTGCCGTGGG CAGCATCCTG AGTGAAGGTG AGGAATCGGC CTCCCCTGAG CTCATCGACC TCTACCAGAA ATTTGGCTTC AAGGTGTTCT CCTTCCCGGGC ACCCAGGCAT GTGGTGACAG CCACCTTTCCC CTACACCACC ATTCTGTCCA TCTGGCTGGC TACCCGGCGT GTCCATCCTG CCTTGGACAC CTACATCAAG GAGCGGAAGC TGTGTGCCTA TCCTCGGTTG GAGATCTACC AGGAAGACCA GATCCATTTC ATGTGCCCAC TGGCASGCA	120 180 240 300 360 420 480 540 600

GAGCAGCCGT GGCTGGGATG ACGGTGACAC CCGCAGCGAG CACAGCTACA GCGAGTCAGG

PCT/US98/04482

	TGCCAGCGGC TCCTCTTTTG AGGAGCTGGA YTTGGAGGGC GAGGGGCCCT TAGGGGAGTC	900
5	ACGGCTGGAC CCTGGGACTK AGCCCCTGGG GACTACCAAG TGGCTCTGGG AGCCCACTGC	960
3	CCCTGAGAAG GGCAAGGAGT AACCCATGGC CTGCACCCTC CCTGCAGTGC AGTTGCTGAG	1020
	GAACTGAGCA GACTCTCCAG CAGACTCTCC AGCCCTCTTC CTCCTTCCTC TGGGGGAGGA	1080
10	GGGGTTCCTG AGGGACCTGA CTTCCCCTGC TCCAGGCCTC TTGCTAAGCC TTCTCCTCAC	1140
	TGCCCTTTAG GCTCCCAGGG CCAGAGGAGC CAGGGGACTAT TTTCTGCAAC CAGCCCCCAG	1200
15	GGCTGCCNCC CCTGTTGTGT CTTTTTTTCA GACTCACAGT GGAGCTTCCA GGACCCAGAA	1260
15	TAAAGCCAAT GATTTACTTG TTTCAAAAAA AAAAWAAAAA AAAAAAAAA AAAAAAAAA AAAAAA	1320
20	(2) INFORMATION FOR SEQ ID NO: 111:	
25	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 1962 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double(D) TOPOLOGY: linear	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:	
30	CGGACCCCTT CCTCCTCCTC NAAGCATGTC CCACCATTGT GGCAGGGGCT GGGGANACAG	60
	TCACCTGATG CGGGGACCAC GGCCACTCCA CCTCGSTGGC GCTGTCAGTG GGCAGCACTG	120
35	GCTGGGCCTG CACTGAGGTC CCTGCTGGGG CAGTTCTTCC AGAATTATCT TCAGAGGGGG	180
	CCTCCAGCTC CCTGGTACCC TCAGGGGGCCC GTGTGGCTGG AAGCAGGGAA GGGGCACCCT	240
40	CGGAGCTTCC TGTCTCCTCG CTCTCTCCTC GAGGGACCCC AGATAGCTCA GGACCACCAG	300
	TTGCCTCCCC CACCTCTCTT GCCTCAACCA GAGTGGAAGG TGATGGGGAT GCTAGGTTCC	360
	TCTCCCTGGG AGTGGGCAGA GTCTCAGIAG CTGGTCCATG GACCCTTGGA GGCCTGGAAG	420
45	CTTCTGACTC TCCATCAGGA AGTGGTGATG CACCAGGCTG CAGGACTGCC CTTGCTGGCG	480
	CCTGGGAGAG TGACTCCTCC TGGGCTGCTG GCTCAGTGGG GAGAGAGGCC TCAGGGCCCG	540
50	GGCTGCTGAG CTCGCTGGGC CATGCCCACA GAGCCTCATC CTCCACCTCC TCCTCTTCTT	
	CTICCTCCTC TITCTCTTCT TCATCTTCAT ATTTCTCTTC TTCCTCCAAT GCCTTACCTT	660
	CCTCTTYTGR AAACCCCGTG GGCGGTACCA TGGATTGTGT TTCAAATTCT AGGAGCGTCC	720
55	TAGGGGCCTC TGCTGGGTCT TCTGGAGTGG AGCTTCCACC TCCTCCGTCC TCCATGATGG	78
	GGATGGAGTA RATGGCCCCA CGGGATTCAC TCTCTGTGGC TTCCTGAGGC AGCTGCAGTI	84
60	CCTCCAGGGT CTCTGTCACT GTGACRATAG CCTCTAGTCC ATCAAAAGCT GGGTTGGAGG	90

	CIGGGTTGGA	GGCCTCAGGG	DAAEALUUTA	GCTGGGCCGA	GTCTCGGAAG	CAGTARACGT	9€0
	TGAAGCGGCT	GTGITTATTG	GGGAAGCCAG	TCTGGTTGGG	GAAGANGAAS	AGAGTCTTGA	1010
5	CACCAGGCAA	ACCCCCCCC	CAGCGCTGGI	CAPTETEDT	GATGGGGTAG	CGCACANTGC	1080
	CATCAGCTAG	CCACCTGGGC	TGCAGTGGTC	CAGGCCACCA	TOCCAGGOTG	CATACAGTTG	1140
10	GCCCGTGGTG	GCAATCTCTG	CACCCCCGCTC	CTGGCAGTAC	GCCCGTGCTT	CCTCCAATGT	12(
10	CAGCTTCTCT	GGAGGGTCAC	CCAGGAACAG	TTCTCCATTT	AGGTCTTCAG	CATAACAGTA	1260
	CACATCATAG	AGGTCATCCG	GGTCCACCAC	ACCATAGTTC	CGGACCCCGG	GGAAGCCATC	1320
15	CATGTCTCCG	TAACAGGCCT	CTCGTG3GGT	CTGGATGGGA	TACCTTTGAC	CTTGAMCTCC	1380
	ACAGCGTCGC	TGCTGTCATC	GATGCCGTGC	TGGACCTCAC	AGCGATAGAT	ACCIGAGTCG	1440
20	TTGGGGCGCA	GCTCGCTCAG	CGCCAGGGGA	GACGTCGGTG	AGCGACGCTG	GGTACGCAGG	1500
20	CAGTGCCACG	CGGAACCGGT	AGGCCTCGTT	CACCTTGACG	CGCACTCCCC	GCGCCACCAG	1560
	CACYTCTGCC	TCCCGGCCCC	GGGACAGGAA	AGTCCACTTG	ACCCGCGGAG	AGCCCAGCAC	1620
25	AGCCCGGCGG	CTCGGCGGTG	SCCGCAGGTA	GTGGACGTGG	CAAGGGATGK	TGAGGGCSCC	1680
	GCCGAGCAAC	GCCYTGCAGT	GGCGCGTCGC	CCGCGATGCG	CACGCGAAAA	GCGCGKTCCT	1740
30	CTGAGCTGTC	TCCTTCCAGA	ACATCTGCIA	AAGCTGCAGG	AGCCTGGGCC	AGGACCAGGG	180
50	CTGCCAGCAG	GGGCAGGAAC	AGCTGGGCCA	TGCTGCAGGC	TACCCAGGGC	TGGGGTTGGG	186
	TCGCGGCACT	GCGAAGTTTG	TCGCCTCCTC	CGGGGGTCTC	CTCCGGGTKC	ACGGCTCAGT	192
35	NCCTGCAGCT	GCAGCTGAGA	CTGCGGCGGA	GACTGCGCGA	GC.		196

40 (2) INFORMATION FOR SEQ ID NO: 112:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1785 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 112:

50 AAGTTTCAGC CAAACTTCGG GCGGCTGAGG CGGCGGCCGA GGAGCGGCGG ACTCSGGGCG 60

CGGGGAGTCG AGGCATTTGC GCCTGGGCTT CGGAGCGTAC CGCAGGGCCT GAGCCTTTGA 120

AGCAGGAGGA GGGGAGGAGA GAGTGGGGCT CCTCTATCGG GACCCCCTCC CCATGTGGAT 180

CTGCCCAGGC GGCGGCGG GCCGAGGAGG CGACCGAGAA GATRCCCGCC CTGCGCCCCG 240

CTCTGCTGTG GGCGCTGCTG GCGCTCTGGC TGTGCTGCGC GACCCCGCGC ATGCATTGCA 300

	GTGTCGAGAT	GGCTATGAAC	CCTGTGTAAA	TGAAGGAATG	TGTGTTACCT	ACCACAATGG	360
	CACAGGATAC	TGCAAATGTC	CAGAAGGCTT	CTTGGGGGAA	TATTGTCAAC	ATCGAGACCC	426
5	CTGTGAGAAG	AACCGCTGCC	AGAATGGTGG	GACTTGTGTG	GCCCAGGCCA	TGCTGGGGAA	480
	AGCCACGTGC	CGATGTGCCT	CAGGGTTTAC	AGGAGAGGAC	TGCCAGTACT	CGACATCTCA	540
10	TCCATGCTTT	GTGTCTCGAG	CTTGCCTGAA	TGGCGGCACA	TGCCATATGC	TCAGCCGGGA	600
10	TACCTATGAG	TGCACCTGTC	AAGTCGGGTT	TACAGGTAAG	GAGTGCCAAT	GGACCGATGC	660
	CTGCCTGTCT	CATCCCTGTG	CAAATGGAAG	TACCTGTACC	ACTGTGGCCA	ACCAGTTCTC	720
15	CTGCAAATGC	CTCACAGGCT	TCACAGGGCA	GAAGTGTGAG	ACTGATGTCA	ATGAGTGTGA	780
	CATTCCAGGA	CACTGCCAGC	ATGGTGGCAC	CTGCCTCAAC	CTGCCTGGTT	CCTACCAGTG	840
20	CCAGTGCCTT	CAGGGCTTCA	CAGGCCAGTA	CTGTGACAGC	CTGTATGTGC	CCTGTGCACC	900
20	CTCGCCTTGT	GTCAATGGAG	GCANCTGTCG	GCAGACTGGT	GACTICACTI	TTGAGTGCAA	960
	CTGCCTTCCA	GAAACAGTGA	GAAGAGGAAC	AGAGCTCTGG	GAAAGAGACA	GGGAAGTCTG	1020
25	GAATGGAAAA	GAACACGATG	AGAATTAGAC	ACTGGAAAAT	ATGTATGTGT	GGTTAATAAA	1080
	GTGCTTTAAA	CTGAATTGAC	ATTAACAGTR	GGTGATCAAC	TTTMCTATGT	GCTTGTGCTT	1140
30	TTGCTTTTGA	TGGAGTAATT	CATTGTTTTC	TTATCCACCT	AAATGCACCC	AGCTGCCCTT	1200
	GATTTTCTCT	GGGCTACTGG	CCTTCACAAC	CCTCTCCCAT	GTACCCTCTC	TGACTTTGGG	1260
	GTAACCCTCC	CCTAACTTAA	AGCTAGAGAA	TTCTGAAACT	GAGGAGGGGA	TCCTCTGTTA	1320
35	ATCAGTGAGC	ACTITTTGAT	GAGCTGATAG	ATGATATATG	AGAGACTATG	CGTGGCACAA	1380
	TACTTTGTTA	CACTCTTCAC	TGATACAAGT	GTTCTAGAGT	GYACACACAA	CCCAAAGATA	144(
40	GAAATAAAAA	GAGGAGCAGT	GTCGGGGAGC	TTGGGGCCTG	GTGTTCCATG	GAGAGGGAGA	1500
••	AAGGAACAAG	CTTGRCCAAT	TCATTCAACT	CCTTATAAAA	ATGATGAGGA	GGCTGAAAAC	1560
	CAAGAATTTT	GATTGGGAAC	AGAATACAAG	CAGCTGAAKC	AGATGAWITA	CTAAGCAACA	1620
45	AAGATCCIGT	TTTTATACAA	ATATCCTTAG	TACAAAAACA	AAARAAGGAA	AACTGTAGGG	1680
	GGGAGTAATG	TGCTAAGTAA	GCAGAATTGC	CTCCAAAAGA	AGTTGTTTCT	AGTTACTCTT	1740
50	TTCCGGGTNG	GGATCTTTAG	NTTCCGGTAT	TGTGGGTATG	GTTCC		1785

(2) INFORMATION FOR SEQ ID NO: 113:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1842 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

55

(xi) SEÇUENCE DESCRIPTION: SEÇ ID NO: 113:

5	GGAGCCTCTC	TTGCAATTIC	TGCCACCGCG	GGCCACCGCG	GOOGCOTGAT	COCGOAGAGG	€ (
S	AAGTCGCGGC	CCTGGAGCGA	TGACCCGCGC	CGGTCCGGGC	3860600036	ADADDETDEE	124
	GOGGCGGGG	CTICTGCTGC	TOPTOPITOR	GOMGOTGTTG	TTAGTCACCG	OGGAGOCGOO	187
10	GAAACCTGCA	GGAGTCTACT	АТЭСААСТЭС	ATACTGGATG	CCTGCTGAAA	AGACAGTACA	240
	AGTCAAAAAT	GTAATGGACA	ACCOUTAGEA	CGCCTATGGI	TTTTACAATA	ACTCTGTGAA	300
15	AACCACAGGC	TGGGGGATCC	TEGAGATCAG	AGCTGECTAT	GGCTCTCAAA	CCCTGAGCAA	360
	TGAGATCATC	ATGTTTGTGG	CT/3GCTTTTT	GGAGGGTTAC	CTCACTGCCC	CACACATGAA	42(
	TGACCACTAC	ACAAACCTCT	ACCCACAGGT	GATCACGAAA	CCTTCCATCA	TGGATAAAGT	48(
20	GCAGGATTTT	ATGGAGAAGC	AAGATAAGTG	GACCCGGAAA	AATATCAAAG	AATACAAGAC	540
	TGATTCATTT	TGBAGACATA	CAGGCTATGT	GATGGCACAA	ATAGATGGCC	TCTATGTAGG	60(
25	AGCAAAGA AG	AGGGCTATAT	TAGAAGEGAC	AAAGCCAATG	ACCCTGTTCC	AGATTCAGTT	660
-5	CCTGAATAGT	GTTGGAGATC	TATTGGATCT	GATTCCCTCA	CTCTCTCCCA	CAAAAAACGG	720
	CAGCCTAAAG	GTTTTTAAGA	GATG3GACAT	GGGACATTGC	TOCGCTCTTA	TCAAGGTTCT	780
30	TCCTGGATTT	GAGAACATCC	TTTTTGCTCA	CTCAAGCTGG	TACACGTATG	CAGCCATGCT	840
	CAGGATATAT	AAACACTG3G	ACTTCAACRT	CATAGATAAA	GATACCAGCA	GTAGTCGCCT	900
35	CTCTTTCAGC	AGTTACCCAG	GGTTTTTGGA	GTCTCTGGAT	GATTTTTACA	TTCTTAGCAG	960
	TGGATTGATA	TTGCTGCAGA	CCACAAACAG	TGTGTTTAAT	AAAACCCTGC	TAAAGCAGTA	1020
	ATACCCGAGA	CTCTCCTGTC	CTGGCAAAGA	GTCCGTGTG3	CCAATATGAT	GGCAGATAGT	1080
40	GGCAAGAGGT	GGGCAGACAT	CTTTTCAAAA	TACAACTCTG	GCACCTATAA	CAATCAATAC	1140
	ATGGTTCTGG	ACCTGAAGAA	AGTAAAGCTG	AACCACAGTC	TIGACAAAGG	CACTCTGTAC	1200
45	ATTGTGGAGC	AAATTCCTAC	ATATGTAGAA	TATTCTGAAC	AAACTGATGT	TCTACGGAAA	126C
	GGATATTGGC	CCTCCTACAA	TGTTCCTTTC	CATGAAAAA	TCTACAACTG	GAGTGGCTAT	1320
	CCACTGTTAG	TTCAGAAGCT	GGGCTTGGAC	TACTITTATG	ATTTAGCTCC	ACGAGCCAAA	1380
50	ATTTTCCGGC	GTGACCAA/3G	GAAAGTGACT	GATACGGCAT	CCATGAAATA	TATCATGCGA	1440
	TACAACAATT	ATAAGAAGGA	TCCTTACAGT	AGAGGTGACC	CCTGTAATAC	CATCTGCTGC	1500
55	CGTGAGGACC	TGAACTCACC	TAACCCAAGT	CCTGGAGGTT	GTTATGACAC	AAAGGTGGCA	1560
	GATATCTACC	TAGCATCTCA	GTACACATCC	TATGCCATAA	GTGGTCCCAC	AGTACAAGGT	1620
	GGCCTCCCTG	TTTTTCGCTG	GGACCGTTTC	AACAAAACTC	TACATCAGGG	CATGSCAGAG	1680
60	CTCTACAACT	سالا تالمششالا تالمس	יייברר ביועבא א א	₹ېنىلىلىلالا ∀اب	አልርታዊጋቸው መልጥልጋቸው	SOMMERANA	1740

- 1 - VV 9F1944 A)

	AGGGAGATGA CGGACTAGAA GACTGTAAAT AAGATACCAA AGGCACTATT TTAGCTATGT	1800
5	TITTTCCCATC AGAATTATGC AATAAAATAT ATTAATTTGT CA	1842
10	(2) INFORMATION FOR SEQ ID NO: 114:	
15	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1960 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear 	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 114:	
20	GAATTOGGCA CGAGOTTOTO CGCGCCCCAG COGDOGGCTG CCAGCTTTTC GGGGCCCCGA	60
20	GTCGCACCCA GCGAAGAGAG CGGGCCCGGG ACAAGCTCGA ACTCCGGCCG CCTCGCCCT.	120
	CCCCGGCTCC GCTCCCTCTG CCCCCTCGGG GTCGGGCCC CACGATGCTG CAGGGCCCTG	180
25	GCTCGCTGCT GCTGCTTTC CTCGCCTCGC ACTGCTTGCCT GGGCTCGGCG CGCGGGCTCT	240
	TCCTCTTTEG CCAGCCCGAC TTCTCCTACA AGCGCAGMAA TTGCAAGCCC ATCCCGGTCA	300
20	ACCTGUAGUT GTBUCACGGO ATCGAATACC ABAACATGCG GCTGCCCAAC CTGCTGGGCC	360
30	ACGAGACCAT GAAGEAGGTG CTGGAGCAGG CCGGCGCTTG GATCCCGCTG GTCATGAAGC	420
	AGTGCCACCC GGACACCAAG AAGTTCCTGT GCTCGCTCTT CGCCCCCGTC TGCCTCGATG	480
35	ACCTAGACGA GACCATCCAG CCATGCCACT CGCTCTGCGT GCAGGTGAAG GACCGCTGCC	540
	CCCCGGTCAT GTCCGCCTTC GGNTTCCCCT GGCCGACAT GCTTGAGTGC GACCGTTTCC	60C
40	CCCAGGACAA CGACCTTTGC ATCCCCCTCG CTAGGAGCGA CCACCTCCTG CCAGCCACCG	660
40	AGGAAGCTCC AAAGGTATGT GAAGCCTGCA AAAATAAAAA TGATGATGAC AACGACATAA	720
	TGGAAACGCT TTGTAAAAAT GATTTTGCAC TGAAAATAAA AGTGAAGGAG ATAACCTACA	780
45	TCAACCGAGA TACCAAAATC ATCCTGGAGA CCAAGAGCAA GACCATTTAC AAGCTGAACG	840
	GTGTGTCCGA AAGGGACCTG AAGAAATCGG TGCTGTGGCT CAAAGACAGC TTGCAGTGCA	900
50	CCTGTGAG3A GAT3AACGAC ATCAACGCGC CCTATCTGGT CATGGGACAG AAACAGGGTC	960
50	GGGAGCTGGT GATCACCTCG GTGAAGCGGT GGCAGAGAGG GCAGAGAGAG TTCAAGCCCA	1020
	TCTCCCGCAG CATCCGCAAG CTGCAGTGCT AGTCCCGGCA TCCTGATGGC TCCGACAGGC	1080
55	CTGCTCCAGA GCACGGCTGA CCATTTCTGC TCCGGGATCT CAGCTCCCGT TCCCCAAGCA	11.40
	CACTCCTAGC TGCTCCAGTC TCAGCCTGGG CAGCTTCCCC CTGCCTTTTG CACGTTTGCA	1200
	TCCCCAGCAT TTCCTGAGTT ATAAGGCCAC AGGAGTGGAT AGCTGTTTTC ACCTAAAGGA	1260
60		

	AAASCCCACC CSAATCTIGI AGAAATATIC AAACTAATAA AATCATGAAT ATTITTATGA	1320
	AGTITAAAAA TAGCTCACTT TAAAGCTAGT TTIGAATAGG DECAACTETG ACTTOFCTCI	1380
5	GETTGETTEE TETTTETT THYSASTOAG CTGATTTTCA CTTCCCACTG AGGTTETCAT	1440
	AACATGCAAA TYGOTTCAAT TTTOIOTGIG GCCCAAACTT GTGGGTCACA AACCCIGTTS	1500
10	AGATAAAGOT GGCTGTTATC TCAACATCTI CATCAGCTCC AGACTGAGAC TCAGTGTCTA	15 6 0
10	AGTOTIACAA CAATTCATCA TITTATACCT TCAATGGGAA CTTAAACTGT TACATGTATC	1620
	ACATTYCAGC TACAATACTT CCATTTATTA GAAGCACATT AACCATYTCT ATAGCATGAT	1680
15	TTCTTCAAGT AAAAGGCAAA AGATATAAAT TYTATAATTG ACTTGAGTAC TTTAAGCCTT	1740
	ATAAAACA TTTCTTACT TAACTTTTAA AOTTTTAA AOTTTTAATA TTOATATTTTOTTAATA	1800
20	TACATAGTAG TITTACCTITA AAAGTIGTAA AAATATTGCT TTAACCAACA CTGTAAATAT	1860
20	TTCAGATAAA CATTATATTC TIGTATATAA ACTTTACATC CTGTTTTACC TAAAAAAAA	1920
	ALALAAAA AAAAACTCG AGGGGGGCCC GGTACCCAAT	1960
25		
	(2) INFORMATION FOR SEQ ID NO: 115:	
• •		
30	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 536 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
25	(D) TOPOLOGY: linear	
35	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 115:	
	GTGCTCAGCC CCCGGGGGAC AGYAGGACGT TTGGGGGGCCT TCTTTCAGCA GCGGACAGCC	60
40	CGATTGGGGA CAATGGCGTC TCTTGGCCAC ATCTTGGTTT TCTGTGTGGG TCTCCTCAC:	120
	ATGGCCAAGG CAGAAAGTCC AAAGGAACAC GACCCGTTCA CTTACGACTA CCAGTCCCTG	180
15	CAGATOGRAG GOOTTOGTOAT COCCOGGGATO OT TYTOATOC TOGGGCATOOT CATCRIGOTY?	240
45	AGCAGAAGAT GCCGGTGCAA GTTCAACCAG CAGCAGAGGA CTGGGGAACC CGATGAAGAG	300
	GAGGERACTT TOTGERAGETO CENTOCHOTO TETECRANCO GEARGEGERA GAARGACTES	360
50	GAGG SATISGA ATTOCOGGODAG GAGTOCODIG GCACOTGACA TOTOCCACGO TOCAGOTSGO	420
	CGCCCACIGO CCCITCOGCO GDICCTTCCC CAGCCCTGCC CCCGCAGACT CCCCCTGCCC	480
55	CGCCCACIGO CCCITCOGCO GDICOTTICC CAGCCCTGCC CCCGCAGACT CCCCCTGCCG CCAAGACITI CAATAAAACG TGCGTTICTC TCGAMAAAAA AAAAAATAAA AAAACI	480 536

(2) INFORMATION FOR SEQ ID NO: 116:

PCT/US98/04482

ATT OF WORKERS

5	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 790 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE LESCFIPTION: SEQ ID NO: 116:	
	GTGGGGAGGG GGCGGAGCAA AGCCGCGCCT CTGGGTGGGC GGGTCGGGCC GTCCAGGTCC	60
10	CTGACTTGAA CCTTCCCGGT CCCCAGCCCT CAACAGGAGG CGCAGAAAAT CTTCAAAGCC	120
	AACCACCCCA TGGACGCAGA AGTTACTAAG GCCAAGCTTC TGGGGTTTGG CTCTGCTCTC	180
15	CTGGACAATG TGGACCCCAA CCCTGAGAAC TTCGTGGGGG CGGGGATCAT CCAGACTAAA	240
	GCCCTGCAGG TGGGCTGTCT GCTTCGGCTG GAGCCCAATG CCCAGGCCCA GATGTACCGG	300
- 0	CTGACCCTGC GCACCAGCAA GGAGCCCGTC TCCCGTCACC TGTGTGAGCT GCTGGCACAN	360
20	AGTTCTGAGC CCTGGACTCT GCCCCG3GG3 ATGTGGCCGG CACTGGGCAG CCCCTTGGAC	420
	TGAGGCAGTT TTGGTGGATG GGGGACCTCC ACTGGTGACA GAGAAGACAC CAGGGTTTGG	480
25	GGGATGCCTG GGACTTTCCT CCGGCCTTTT GTATTTTTAT TTTTGTTCAT CTGCTGCTGT	540
	TTACATTCTG GGGGGTTAGG GGGAGTCCCC CTCCCTCCCT TTCCCCCCCA AGCACAGAGG	600
20	GGAGAGGGGC CAGGGAAGTG GATGTCTCCT CCCCTCCCAC CCCACCCTGT TGTAGCCCCT	660
30	CCTACCCCCT CCCCATCCAG GGGCTGTGTA TTATTGTGAG CGAATAAACA GAGAGACGTT	720
	AACAGCCCCA TGTCTGTGTC CATCACCCAN TGNTAGGTAG TCAAAGAAGT GGGGTGAGGG	78(
35	CATGCAGAGT	790
40 45	(2) INFORMATION FOR SEQ ID NO: 117: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 776 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 117:	
50	CAGCGCTGGA AGCAGCTGAG CCTGTGAGGG GTGGGGAGGGG GGCGGAGCAA AGCCGCGCCT	6C
	CTGGGTGGGC GGGTCGGGCC GTCCAGGTCC CTGACTTGAA CCTTCCCGGT CCCCAGCCCT	120
<i></i>	CAACAGGAGG CGCAGAAAAT CTTCAAAGCC AACCACCCCA T3GACGCAGA AGTTACTAAG	180
55	GCCAAGCTTC TGGGGTTTGG CTCTGCTCTC CTGGACAATG TGGACCCCAA CCCTGAGAAC	24(
	TTCGTGGGGG CGGGGATCAT CCAGACTAAA GCCCTGCAGG TGGGTTGTCT GTTTCGGCTG	300
60	GAGCCCAATG CCCAGGCCCA GATGTACCGG CTGACCCTGC GCACCAGCAA GGAGCCCGTC	360

	iscasione seriose serioseveno nelicianos celebratel estacogodo	420
5	ATGT99C09G CACTG9GCAG CCCTTGGAC TGAGGCAGTT TIGGTGGATG GG9GACCTCC	480
J	ACTGGTGACA GAGAAGACAC CAGGGTTTGG GGGATGCCTG GGACTTTCCT CCGGCCTTTT	540
	GTATTTTAT TTTTGTTCAI CTGCT3CTGT TTACATTCTG GGGGGTTAGG GGGAGTCCCC	600
10	CTCCCTCCCT TTCCCCCCCA AGCACAGAGG GGAGAGGGGC CAGGGAAGTG GATGTCTCCT	66(
	CCCCTCCCAC CCCACCCTGT TGTAGCCCCT CCTACCCCCT CCCCATCCAG GGGCTGTGTA	720
15	TTATTGTGAG CGAATAAACA GAGAGACGCN TAAAAAAAAA AAAAAAAAT TGAGGG	77€
20	(2) INFORMATION FOR SEQ ID NO: 118: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 453 base pairs	
25	(B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 118:	
30	GGTTCTGACA CCAGATGTTC TCTGCTCCTG GTTAATGTCA GTGAGGGCTG GAAGTTGAAT	60
	AAATGAGAAC AGGAGTGGTC TGGGCCCATG TAAATGATCC TCCCTTGAAA GGAGGAACAG	126
	CTTTCATCAT TTGTTCCAGC TAAGCCTTGC ATGCATTATA GATCTGGTGC TAAGCAGTGG	180
35	GAAAGATCTC ATAAGTAATG TTTTATGTTC TTTTCKGTCTC TCYTCTTCKG TTGTTCTTGG	240
	CTTGTGGGTT GTGTTTGKGG TTGTTAACTG GAAAATTGCT ATAAGCCAGT TGTCYCKAAK	300
40	TITWAAAAAC GAATTAGAAA AACCATAAAA TCYTCTGGCC YATGCACATK GTCCCYGTTT	360
	TGTGAAAACA TTAAAGGGTA AATAAAAAGG AAGGAGAACA GTCAATAATG TGCATCAAAT	420
45	ATATTCTGAG TTCTAGAGAA ATTAATGACC AAG	45 3
	(2) INFORMATION FOR SEQ ID NO: 119:	
50	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2016 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
55	(xi) SEQUENCE DESCRIPTION: SEO ID NO: 119:	
	AGGCTGTTCA CAGGCACCCC GAGACAGCGT CCCCCCTCTG GGCGCACTGG ATTTGACGTT	60
60	GCAGGACGCG CGGCTGGAAC CCCCAGGCCC CGCTGCTCAC AGACCGGGAC TCCGCCTCCG	330

	GTTCCCGAGG	GCCTGGCGAG	GCGCTGCGGG	ANCCCAACAG	GATGCCTTCC	GTGCCTTCCA	180
_	TCAAGATCTC	AATTTTGTGC	GCAATTCCTA	CAGCCCCTGT	TGATTGGAGA	GCTGGCTCCG	240
5	GAAGAACCCA	GCCAKGATGG	ACCCCTGAAT	GCGCATGGTC	GAGGACTTCC	GAGCCCTGCA	300
	CCAGGCAGCC	GAGGACATGA	AGCTGTTTGA	TGCCAGTCCC	ACCTTCTTTG	CTTTCCTACT	360
10	GGGCCACATC	CTGGCCATG3	AGGTGCTGGC	CTGGCTCCTT	ATCTACCTCC	TGGGTCCTGG	420
	CTGGGTGCCC	AGTGCCCTGG	NCCGCCTTCA	TCCTGGCCAT	CTCTCAGGCT	CAGTCCTGGT	4 80
	GTCTGCAGCA	TGACCTGGGC	CATGCTCCAT	CTTCAAGAAG	TCCTGGTGGA	ACCACGTGGC	540
15	CCAGAAGTTC	GTGATGGGGC	AGCTAAAGGG	CTTCTCCGCC	CACTGGTGGA	ACTTCCGCCA	600
	CTTCCAGCAC	CACGCCAAGC	CCAACATCTT	CCACAAAGAC	CCAGACGTGA	CGGTGGCGCC	660
20	CGTYTTCCTC	CTGGGGGAGT	CATCCGTCGA	GTATGGNCAA	GAAGAAACGC	AGATACCTAC	720
	CCTACAACCA	GCAGCACCTG	TACTTCTTCC	TGATCGGCCC	GCCGCTGCTC	ACCCTGGTGA	780
	ACTTTGAAGT	GGAAAATCTG	GCGTACATGC	TGGTGTGCAT	GCAGTGGGCG	GATTIGCTCT	8 4 C
25	GGGCCGCCAG	CTTCTATGCC	CGCTTCTTCT	TATCCTACCT	CCCCTTCTAC	GGCGTCCCTG	900
	GGGTGCTGCT	CTTCTTTGTT	GCTGTCAGGT	ATGGCAGGGA	GTGGCGAGGT	CACACACAGG	96(
30	CGACAGGTGA	CCCCCACTGC	AGCCCCCCAC	CAGAGCTTCC	CTTTTCCCGT	CTGCAGAATG	1020
	GGGCCAGTGG	TACTGCCTCC	CTGGCTTGCT	GGTGGAATCA	. CATAAACACA	AGYTTCAGGA	1080
0.5	GCCCAGGGTC	GGTGGGTTTA	GGGAGCGTGG	CCTGGCTTGT	· AAGTGGCCCG	GTGGGTGTCG	114(
35	GAGCTGCTCT	GGACTCAGCC	TCACAGTGGA	CACTGCTCCA	TTCAGATTCT	TTAAACACTG	1200
	GCAAGGGGGC	GATGGCCACA	ATCCTATTGT	ACAGATAAGG	; AAGTCAAGGC	CAYTTGGGGA	1260
40	CAGYTGCTCT	TCCAGCCTCC	ACTCAGGGTG	CCTTAAGTGC	G TGAGCTGGAC	CTAGGGCAGT	1320
	GCCGAGCYTC	CCCACAGGGT	CCTGGAAAGC	CACTGGTTCC	TGTGGATCAC	CACAGATGAAC	1380
	CACATCCCCA	AGGAGATCGG	CCACGAGAAG	CACCGGGACT	GGGTCAGCTC	TCAGCTGGCA	1440
45	GCCACCTGCA	ACGTGGAGCC	CTCACTTITC	ACCAACTGGT	TCAGCGGGCA	A CCTCAACTTC	1500
	CAGATCGAGO	ACCACCTCT1	CCCCAGGATG	CCGAGACACA	A ACTACAGCCC	GGTGGCCCCG	1560
50	CTGGTCAAGT	CGCTGTGTGC	CAAGCACGGC	CTCAGCTACC	CODRAGETAL E	TTCCTCACCG	1620
	CGCTGGTGGA	CATCGTCAGO	TCCCTGAAGA	AGTCTGGTG	A CATCTGGCT	GACGCCTACC	1680
	TCCATCAGTC	AAGGCAACAC	CCAGGCGGG	AGAGAAGGG	TCAGGGCAC	C AGCAACCAAG	1740
55	CCAGCCCCC	GCGGGATCGA	A TACCCCCAMO	CCTCCACTG	G CCAGCCTGG	G GGTGCCCTGC	1800
	CTGCCCTCCT	r ggtactgtt	TCTTCCCCT	GGCCCCCTC	A CATGTGTAT	I CAGCAGCCCT	1860
60	ATGGCCTTGC	G CTCTGGGCCT	r gatgggacac	GGGTAGAGG	G AAGGTGAGC	A TAGCACATTT	1920

PCT/US98/04482

	TCCLAGAGE AGAATACAA TETELEAGA GOODELAAD TETELEAGA TETELE	1980
E	AAAAAAAAA AAAAAAANCT CGAGGGGGGG CCCCGG	201€
5		
	(2) INFORMATION FOR SEQ ID NO: 126:	
10	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 2136 base pairs(B) TYPE: nucleic acid	
15	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 120:	
20	GGGGACGGAG CCGCTGTCAA CTCTCCAACT CAGCTCAGCT	60
20	GCCGCCAGAT TCTGGAGGCG AAGAACGCAA AGCTGAGAAC ATGGACGTTA ATATCGCCCC	120
	ACTCCGCGCC TGGGACGATT TCTTCCCGGG TTCCGATCGC TTTGCCCGGC CGGACTTCAG	180
25	GGACATTTCC AAATGGAACA ACCGCGTAGT GAGCAACCTG CTCTATTACC AGACCAACTA	240
	CCTGGTGGTG GCTGCCATGA TGATTTCCAT TGTGGGGTTT CTGAGTCCCT TCAACATGAT	300
30	CCTGGGAGGA ATCGTGGTGG TGCTGGTGTT CACAGGGTTT GTGTGGGCAG CCCACAATAA	360
20	AGACGTCCTT CGCCGGATGA AGAAGCGCTA CCCCACGACG TTCGTTATGG TGGTCATGTT	420
	GGCGAGCTAT TTCCTTATCT CCATGTTTGG AGGAGTCATG GTCTTTGTGT TTGGCATTAC	480
35	TTTTCCTTTG CTGTTGATGT TTATCCATGC ATCGTTGAGA CTTCGGAACC TCAAGAACAA	540
	ACTGGAGAAT AAAATGGAAG GAATAGGTTT GAAGAGGACA CCGATGGGCA TTGTCCTGGA	600
40	TGCCCTAGAA CAGCAGGAAG AAGGCATCAA CAGACTCACT GACTATATCA GCAAAGTGAA	660
	GGAATAAACA TAACTTACCT GAGCTAGGGT TGCAGCAGAA ATTGAGTTGC AGCTTGCCCT	720
	TGTCCAGACC TATKITCTGC TTGCGTTTTT GAAACAGGAG GTGCACGTAC CACCCAATTA	780
45	TCTATGGCAG CATGCATGTA TAGGCCGAAC TATTATCAGC TCTGATGTTT CAGAGAGAAG	840
	ACCTCAGAAA CCGAAAGAAA ACCACCACCC TCCTATTGTG TCTGAAGTTT CACGTGTGTT	
50	TATGAAATCT AATGGGAAAT GGATCACACG ATTTCTTTAA GGGAATTAAA AAAAATAAAA	960
	GAATTACGGC TTYTACAGCA ACAATACGAT TATCTTATAG GAAAAAAAA ATCATTGTAA	
	AGTATCAAGA CAATACGAGT AAATGAAAAG GCTGTTAAAG TAGATGACAT CATGTGTTAG	1080
55	CCTGTTCCTA ATCCCCTAGA ATTGTAATGT GTGGGATATA AATTAGTTTT TATTATTCTC	
	TTAAAAATCA AAGATGATCT CTATCACTTT GCCACCTGTT TGATGTGCAG TGGAAACTGG	1200
60	TTAAGCCAGT TGTTCATACT TCSTTTACAA ATATAAAGAT AGCTGTTTAG GATATTTTGI	1200

	TACATIFITG TAAATTITTG AAATGCTAGT AATGTGTFTT CACCAGCAAG TATTTGTTGC	1320
	AAACTTAATG TCATTTTCCT TAAGATGGTT ACAGCTATGT AACCTGTATT ATTCTGGACG	1380
5	GACTTATTAA AATACAAACA GACAAAAAAT AAAACAAAAC	1440
	ACATTTTTG TIGTTACAGT GAAAAAAATG GTCCAAGAAA ATGTTTGCCA TTTTTGCATT	1500
10	CTTTCGTTTT TAACTGAAC ATTTAGAAAG AAGGAAATGA TTTTAGTTAT TATTAATTCC	1560
10	TTAGGGGCAC AAGAAGGACA ATAATAGCTG ATCTTTTGAA ATTTGAAAAA CGTCTTTAGA	1620
	TGACCAAGCA AAAAGACTTT AAAAAATGCT AATGAAAATG GAATGCAGCT ACTGCAGCTA	1680
15	ATAAAAATT TTAGATAGCA ATTGTTACAA CCATATGCCT TTATAGCTAG ACATTAGAAT	1740
	TATGATAGCA TGAGTTTATA CATTCTATTA TTTTTCCTCC CTTTCTCATG TTTTTATAAA	1800
20	TAGGTAATAA AAAATGTTTT GCCTGCCAAT TGAATGATTT CGTAGCTGAA GTAGAAACAT	1860
20	TIAGSTTTCT GTAGCATTAA ATTGTGAAGA CAACTGGAGT GGTACTTACT GAAGAAACTC	1920
	TCTGTATGTC CTAGAATAAG AAGCAATGAT GTGCTGCTTC TGATTTTTCT TGCATTTTAA	1980
25	ATTCTCAGCC AACCTACAGC CATGATCTTI AGCACAGTGA TATCACCATG ACTTCACAGA	2040
	CATGGTCTAG AATCTGTACC CTTACCCACA TATGAAGAAT AAAATTGATT AAAGGTTAAA	2100
30	AAAAAAWAA AAAAAMWAGG GGGCCCGGT WCCCAG	2136
30		
35	(2) INFORMATION FOR SEQ ID NO: 121:	
55	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 219 base pairs	
40	(B) TYPE: nucleic acid(C) STRANDEDNESS: doubl∈(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 121:	
15	GCCCTAGTAT CTGGGCAGCT GTGCATGGAG ATAGCCAGAG GAAACATTTT TTTTCTTAAT	60
45	GRATTGGTGA CCACATTTTG TTGTTCTTGC CTCCTATTAT CCGTGCSCTA TTTGCATSCT	120
	GGTTTCTTCT ACAGTAGTTT ATGTAAATGT TGTTTTGTCC TTGTCGTTCT CAGTAGAATT	180
50	GGTTCTGTAA ACGAAACCTG GTCCTGTAAT TTCAGTATA	219

55 (2) INFORMATION FOR SEQ ID NO: 122:

(i) SEÇUENCE CHARACTERISTICS:

(A) LENGTH: 1686 base pairs

(B) TYPE: nucleic acid

60 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 122:

5	GCTGGAGATT	CACATTTTAC	CTGATTGCCT	TCATTGCCGG	CATGGCCGTC	ATTGTGGATA	60
	AACCCTGGTT	CTATGACATG	AAGAAAGTTT	GGGAGGGATA	TCCCATACAG	AGCACTATCC	120
10	CTTCCCAGTA	TTGGTACTAC	ATGATTGAAC	TTTCCTTCTA	CTGGTCCCTG	CTCTTCAGCA	180
10	TTGCCTCTGA	TGTCAAGCGA	AAGGATTTCA	AGGAACAGAT	CATCCACCAT	GTGRCCACCA	240
	TCATTCTCAT	CAGCTTTTCC	TGGTTTGCCA	ATTACATCCG	AGCTGGGACT	CTAATCATGG	300
15	CTCTGCATGA	CTCTTCCGAT	TACCTGCTGG	AGTCAGCCAA	GATGTTTAAC	TACGCGGGAT	360
	GGAAGAACAC	CTGCAACAAC	ATCTTCATCG	TCTTCGCCAT	TGTTTTTATC	ATCACCCGAC	420
20	TGGTCATCCT	GCCCTTCTGG	ATCCTGCATT	GCACCCTGGT	GTACCCACTG	GAGCTCTATC	480
20	CTGCCTTCTT	TGGSTATTAC	TTCTTCAATT	CCATGATGGG	AGTTCTACAG	CTGCTGCATA	540
	TCTTCTGGGC	CTACCTCATT	TTGCGCATGG	CCCACAAGTT	CATAACTGGG	AAAGCTGGTA	600
25	GAAGATGAAC	GCAWGCRCGG	GNAAGAAACA	GAGAGCTCAG	AGGGGGAGGA	GGCTGCAGCT	660
	GGGGGAGGAG	CAAAGAGCCG	GCCCCTAGCC	AATGGCCACC	CCATCCTCAA	TAACAACCAT	720
30	CGTAAGAATG	ACTGAACCAT	TATTCCAGCT	GCCTCCCAGA	TTAATGCATA	AAGCCAAGGA	780
30	ACTACCCYGC	TCCCTGCGCT	ATAGGGTCAC	TTTAAGCTCT	GGGGAAAAAG	GAGAAAGTGA	840
	GAGGAGAGTT	CTCTGCATCC	TCCCTCCTTG	CTTGTCACCC	AGTTGCCTTT	AAACCAAATT	900
35	CTAACCAGCC	TATCCCCAGG	TAGGGGGACG	TTGGTTATAT	TCTGTTAGAG	GGGGACGGTC	960
	GTATTTTCCT	CCCTACCCGC	CAAGTCATCC	TTTCTACTGC	TTTTGAGGCC	CTCCCTCAGC	1020
40	TCTCTGTGGG	TAGGGGTTAC	AATTCACATT	CCTTATTCTG	AGAATTTGGC	CCCAGCTGTT	1080
,,,	TGCCTTTGAC	TCCCTGACCT	CCAGAGCCAG	GGTTGTGCCT	TATTGTCCCA	TCTGTGGGCC	1140
	TCATTCTGCC	AAAGCTGGAC	CAAGGUTAAC	CTTTCTAAGC	TCCCTAACTT	G/GCCAGAAA	1200
45	CCAAAGCTGA	GCTTTTAACT	TTCTCCCTCT	ATGACACAAA	TGAATTGAGG	GTAGGAGGAG	1260
	GGTGCACATA	ACCCTTACCC	TACCTCTGCC	AAAAAGTGGG	GGCTGTACTG	GGGACTGCTC	1320
50	GGATGATCTT	TCTTAGTGCT	ACTTCTTTCA	GCTGTCCCTG	TAGCGACAGG	TUTAAGATCT	1380
	GACTGCCTCC	TOCTFTCTCT	GGCCTCTTCC	CCCTTCCCTC	TTCTCTTCAG	CTAGGCTAGC	1440
	TGGTTTGGAG	TAGAATGGCA	ACTAATTCTA	TTTATTTTA .	PTATAAATTA :	TGGGGTTTTG	1500
55	GTTTTAAAGC	CAGAATTACG	GCTAGCACCT	AGCATTICAC	CAGAGGGACC	ATTTTAGACC	1560
	AAAATGTACT	GTTAATGGGT	ATTTTTTTT .	A AATTAAAAGA	AAAATAAATT A	TAAATTATAA A	1620
60	AAAACATGGC	AATAAGTGTC	: AGACTATTAC	GAATTGAGAI	A GGGGGATCAA	A CTAAATAAAC	1680

1211

GAAGAG	1686
(2) INFORMATION FOR SEQ ID NO: 123:	
(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 1211 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double(D) TOPOLOGY: linear	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 123:	
CAGCCTGTGC CAGACGAGGA GGTGATTGAG CTGTATGGGG GTACCCAGCA CATCCCACTA	60
TACCAGATGA GTGGCTTCTA TGGCAAGGGT CCCTCCATTA AGCAGTTCAT GGACATCTTC	120
TOGOTACOGG AGATGGOTOT GOTGTOCTGT GTGGTGGACT ACTTTCTGGG CCACAGCCTG	180
GAGTTTGACC AAACATCTCT ACAAGGACGT GACGGACGCC ATCCGAGACG TGCATGTGAA	240
GGGCCTCATG TACCAGTGGA TCGAGCAGGA CATGGAGAAG TACATCCTGA GAGGGGATGA	300
GACGTTTGCT GTCCTGAGCC GCCTGGTGGC CCATGGGAAA CAGCTGTTCC TCATCACCAA	360
CAGTCCTTTC AGCTTCGTAG ACAAGGGGAT GCGGCACATG GTGGGTCCCG ATTGGCGCCA	420
CTCTTCGATG TGGTCATTGT CCAGGCAGAC AAGCCCAGCT TCTTCACTGA CCGGCGCAAC	480
TTTCAGAAAA CTCGATGAGA AGGGCTCACT TCAGTGGGAC CGGATCACCC GCTTGGAAAA	540
GGGCAAGATC TATCGGCAGG GAAACCTGTT TGACTTCTTA CGCTTGACGG AATGGCGTGG	600
CCCCCGCGTG CTCTACTTCG GGGACCACCT CTATAGTGAT CTGGCGGATC TCATGCTGCG	660
GCACGGCTGG CGCACAGGCG CCATCATCCC CGAGCTGGAG CGTGAGATCC GCATCATCAA	720
CACGGAGCAG TACATGCACT CGCTGACGTG GCAGCAGGCG CTCACGGGGC TGCTGGAGCG	780
CATGCAGACC TATCAGGACG CGGAGTCGAG GCAGGTGCTG GCTGCCTGGA TGAAAGAGCG	840
GCAGGAGCTG AGGTGCATCA CCAAGGCCCT GTTCAATGCG CAGTTCGGCA GCATCTTCCG	900
CACCTTCCAC AACCCCACCT ACTTCTCAAG GCGCCTCGTG CGCTTCTCTG ACCTCTACAT	960
GGCCTCCCTC AGCTGCCTGC TCAACTACCG CGTGGACTTC ACCTTCTACC CACGCCGTAC	1020
GCCGCTGCAG CACGAGGCAC CCCTCTGGAT GGACCAGCTT CTGCACCGGC TGCATGAAGA	1080
CCCCCTTCCT TGGTGACATG GCCCACATCC GCTGAGGGCA CCTTTATTGT CTGGGACAGG	1140
	(2) INFORMATION FOR SEQ ID NO: 123: (1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1211 base pairs (B) TYPE: nucleic acid (C) STRANDENESS: double (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 123: CAGCCTGTGC CAGACGAGGA GGIGATTGAG CTGTATGGGG GTACCCAGCA CATCCCACTA TACCAGATGA GTGGCTTCTA TGGCAAGGGT CCCTCCATTA AGCAGTTCAT GGACATCTTC TCGCTACCGG AGATGGCTCT GTGTCCTGT GTGGTGGACT ACTTTCTGGG CCACAGCCTG GAGTTTGACC AAACATCTCT ACAAGGACGT GACGGACGCC ATCCGAGAGC TGCATGTGAA GGGCCTCATG TACCAGTGGA TCGAGCAGGA CATGGAGAAG TACATCCTGA GAGGGGATGA GACGTTTGCT GTCCTGAGCC GCCTGGTGGC CCATGGGAAA CAGCTGTTCC TCATCACCAA CAGTCCTTTC AGCTTCGTAG ACAAGGGGAT GCGGCACATG GTGGGTCCCG ATTGGCGCCA CTCTTCGATG TGGTCATTGT CCAGGCAGAC AAGCCCAGCT TCTTCACTGA CCGGGGCACA CTCTTCGATG TGGTCATTGT CCAGGCAGAC AAGCCCAGCT TCTTCACTGA CCGGGGCACA CTCTTCGATG TGGTCATTGT CCAGGCAGAC AAGCCCAGCT TCTTCACTGA CCGGGGCACA CTCTCGATG TGGTCATTGT CCAGGCAGAC AAGCCCAGCT TCTTCACTGA CCGGGGCACA CCCCCGCGTG CTCTACTTCG GGGACCACCT CTATAGTGAT CTGGCGGAT TCATGCTGGG GCACGGCTGG CGCACAGGGG CAACACCTC TATAGTGAT CTGGCGGATC TCATGCTGGG GCACGGCTGG CGCACAGGGG CAACACCTC TATAGTGAT CTGGCGGATC TCATGCTGGG GCACGGCTGG CGCACAGGGG CAACACCTC CTATAGTGAT CTGGCGGATC TCATGCTGGG GCACGGCTGG CGCACAGGGG CAACACCTC CTATAGTGAT CTGGCGGATC TCATGCTGGC GCACGGCTGG CGCACAGGGG CAACACCTC CTATAGTGAT CTGGCGGATC TCATGCTGAC CACGGAGCAG TACATGCACT CGCTGACGTG GCAGCAGGGC CTCACGGGG CTCATCACA CACGGAGCAG TACATGCACT CGCTGACGTG GCAGCAGGGC CTCACGGGG CAACACCTC CATGCAGACC TATCAGGACG CGGAGTCGAG GCAGGTGCTG GCTGCTGGA GCATCTTCCG CACGTTCCAC AACCCCACCT ACTTCTCAAG GCGCCTCGTG CGCTTCTGG ACCTTCCAC GCCCTCCCTA ACCTCCTCC ACCTTCCAACTTC CCTGGACTTC ACCTTCTTCT ACCCTTCCTACCTTCC CACCTTCCAC AACCCCACCT ACTTCTCAAG GCGCCTCGTG CGCTTCTTG ACCTCTACAT GGCCTCCCTC AGCTGCCTGC TCAACTACCC CGTGGACTTC ACCTTCTACC CACGCCTTAC GCCCTCCCTA ACCTGCTGC TCAACTACCC CGTGGACTTC ACCTTCTACC CACGCCTTAC GCCCTCCCAC ACCTGCTCC TCAACTACCC CGTGGACTTC ACCTTCTACC CACGCCTTAC

CCCTCAGCCC CTCCTGCCCC ATCCACCCAG ACAAGCAATA AAAGTGGTCT CCTCCCTGAA

The Conviction of the court of

55

A AAAAAAAA

(2) INFORMATION FOR SEQ ID NO: 124:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1804 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 124. 10 CGCACCTATG GGCTCGCTAC CAGGACATGC GGAGACTGGT GCACGACCTC CTGCCCCCCG 60 AGGTCTGCAG TCTCCTGAAC CCAGCAGCCA TCTACGCCAA CAACGAGATC AGCCTGCGTG 120 ACGTTGAGGT CTACGGCTTT GACTACGACT ACACCCTGGC CCAGTATGCA GACGCACTGC 180 15 ACCCCGAGAT CTTCAGTACC GCCCGTGACA TCCTGATCGA GCACTACAAG TACCCAGAAG 240 GGATTCGGAA GTATGACIAC AACCCCAGCT TIGCCATCCG TGGCCTCCAC TATGACATTC 300 20 AGAAGAGCCT TCTGATGAAG ATTGACGCCT TCCACTACGT GCAGCTGGGG ACAGCCTACA 360 GGGGCCTCCA GCCTGTGCCA GACGAGGAGG TGATTGAGCT GTATGGGGGT ACCCAGCACA 420 TCCCACTATA CCAGATGAGT GGCTTCTATG GCAAGGGTCC CTCCATTAAG CAGTTCATGG 480 25 ACATCTTCTC GCTACCGGAG ATGGCTCTGC TGTCCTGTGT GGTGGACTAC TTTCTGGGCC 540 ACAGCCTGGN AGTTTGACCA AGCACATCTC TACAAGGACG TGACGGACGC CATCCGAGAC 600 30 GTGCATGTGA AGGGCCTCAT GTACCAGTGG ATCGAGCAGG ACATGGAGAA GTACATCCTG 660 AGAGGGGATG AGACGTTTGC TGTCC'IGAGC CGCCTGGTGG CCCATGGGAA ACAGCTGTTC CTCATCACCA ACAGTCCTTT CAGCTTCGTA GACAAGGGGA TGCGGCACAT GGTGGGTCCC 780 35 GATTGGCGCC ACTCTTCGAT GTGGTCATTG TCCAGGCAGA CAAGICCAGC TTCTTCACTG 840 ACCGGCGCAA GCTTTTCAGA AAACTCGATG AGAAGGGCTC ACTTCAGTGG GACCGGATCA 900 40 960 CCCGCTIGGA AAAGGGCAAG ATCTATCGGC AGGGAAACCT GTTTGACTTC TTACGCTIGA CGBAATGGCG TGGCCCCGGC GTGCTCTACT TCGGGGACCA CCTCTATAGT GATCTGGCGG 1020 ATCTCATGCT GCGGCACGGC TGGCGCACAG GCGCCATCAT CCCCGAGCTG GAGCGTGAGA 1080 45 TCCGCATCAT CAACACGGAG CAGTACATGC ACTCGCTGAC GTGGCAGGAG GCGCTCACGG 1140 GGCTGCTGGA GCGCATGCAG ACCTATCAGG ACGCGGAGTC GAGGCAGGTG CTGGCTGCCT 1200 50 GGATGAAAGA GCGGCAGGAG CTGAGGTGCA TCACCAAGGC CCTGTTCAAT GCGCAGTTCG 1260 GCAGCATCTT CCGCACCTTC CACAACCCCA CCTACTTCTC AAAGGCGCCT CGTGCGCTTC 1320 TOTGACCTOT ACATGGCCTC COTCAGCTGC CTGCTCAACT ACCGCGTGGA CTTCACCTTC 1380 55 TACCCACGCC GTACGCCGCT GCAGCACGAG GCACCCCTCT GGATGGACCA GCTCTGCACC 1440 1500 GGCTGCATGA AGACCCCCTT CCTTGGTGAC ATGGCCCACA TCCGCTGAGG GCACCTTTAT

WO 98/39446 PCT/US98/04482

269

	TGTCTGGGAC	AGGCCCTCAG	CCCCTCCTGC	CCCATCCACC	CAGACAAGCA	ATAAAAGTGG	1560
	TCTCCTCCCT	GTGCATGCTT	CTGCTTTCAG	CCCCAGCCTC	GTCACTTGAC	TGTGAGGATC	1620
5	CTCTGGGTGT	CAGGGAAGTC	CTCCTCCAGC	AGTGAGTCAT	CGAAGGGTTC	ACAAAAGGTG	1680
	TCGCTGCCAA	AGACAGGGTT	GGGGACAGAG	ACCAGGGTGG	GETTGGTCCC	TTCTTGCCAC	1740
10	GGTGAGAAGT	CGTCGTCAGC	CGGACGCGTG	GGTCGACCCG	GGAATTCCGG	ACCGGTACCT	1800
10	GCAG						1804

15

20

(2) INFORMATION FOR SEQ ID NO: 125:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1282 base pairs(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 125:

25 60 CCGCAGGNCA GCGACGCGAC TCTGGTGCGG GCCGTCTTCT TCCCCCCGAG CTGGGCGTGC GCGGCCGCAA TGAACTGGGA GCTGCTGCTG TGGCTGCTGG TGCTGTGCGC GCTGCTCCTG 120 CTCTTGGTGC AGCTGCTGCG CTTCCTGAGG GCTGACGCG ACCTGACGCT ACTATGGGCC 180 30 GAGTGGCAGG GACGACGCCC AGAATGGGAG CTGACTGATA TGGTGGTGTG GGTGACTGGA 240 300 GCCTCGAGTG GAATTGGTGA GGAGCTGGCT TACCAGTTGT CTAAACTAGG AGTTTCTCTT 35 GTGCTGTCAG CCAGAAGAGT GCATGAGCTG GAAAGGGTGA AAAGAAGATG CCTAGAGAAT 360 GGCAATTTAA AAGAAAAAGA TATACTTGTT TTGCCCCTTG ACCTGACCGA CACTGGTTCC 420 480 CATGAAGCGG CTACCAAAGC TGTTCTCCAG GAGTTTGGTA GAATCGACAT TCTGGTCAAC 40 540 AATGGTGGAA TGTCCCAGCG TTCTCTGTGC ATGGATACCA GCTTGGATGT CTACAGAAAG CTAATAGAGC TTAACTACTT AGGGACGGTG TCCTTGACAA AATGTGTTCT GCCTCACATG 600 45 ATCGAGAGGA AGCAAGGAAA GATTGTTACT GTGAATAGCA TCCTGGGTAT CATATCTGTA 660 720 CCTCTTTCCA TTGGATACTG TGCTAGCAAG CATGCTCTCC GGGGTTTTTT TAATGGCCTT 780 CGAACAGAAC TTGCCACATA CCCAGGTATA ATAGTTTCTA ACATTTGCCC AGGACCTGTG 50 840 CAATCAAATA TTGTGGAGAA TTCCCTAGCT GGAGAAGTCA CAAAGACTAT AGGCAATAAT GGAGACCAGT CCCACAAGAT GACAACCAGT CGTTGTGTGC GGCTGATGTT AATCAGCATG 900 55 GCCAATGATT TGAAAGAAGT TTGGATCTCA GAACAACCTT TCTTGTTAGT AACATATTTG 960 1020 TGGCAATACA TGCCAACCTG GGCCTGGTGG ATAACCAACA AGATGGGGAA GAAAAGGATT 1080 GAGAACTTTA AGAGTGGTGT GGATGCAGAC TCTTCTTATT TTAAAATCTT TAAGACAAAA 60

WO 98/39446

	CATGACTGAA AAGAGCAYCT GTACTTTTCA AGCCACTGGA GGGARAAATG GAAAACATGA	1140
5	AAACAGCAAT CTTCTTATGC TTCTGAATAA TCAAAGACTA ATTTGTGRTT TTACTTTTTA	1200
	ATAGATATGA CTTTGCTTCC AACATGGAAT GAAATAAAA ATAAATAATA AAAGATTGCC	1260
	atggaaaaa aaaaginggg an	1282
10		
	(2) TAYONARYON TOO ONE TO US	
15	(2) INFORMATION FOR SEQ ID NO: 126:	
20	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1296 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 126:	
	GGCAGAGCTT AGAGTGTGGA AAAGGCAACC AGGTTGGCCG TAAGTGCCTG CTGGAATGCG	60
25	TGTGCCTCCA CASSSFTCTG GGCATCCGGA CTGATAACCA GCCGGCCAGA CTGAGGGATG	120
	GAAGGCACTG AGATGGGGGC CCGTCCAGGC GGACACCCGC AGAAATGGAG CTTTCTGTGG	180
30	TOTOTTGCAC TOTASCIGOO TOTTGCCCTC TOTGTGTCTC TOTTTCTTGG TOTCTCCCTC	240
50	TCTCCTCCTC AGCCTGGTCT TTCTCTTTGG TGCACACTTA GTTATTGTTG TGAGCAATGG	300
	AAGTTCAAAG GAACTCCCTC TCCAGCTCTT CTGAATCTTG GGACACAGCC TAAAAAGGAC	360
35	AAAAAGTTAG AAGACAGCAT AGCAACTCAG CTCAGGGRGC TACCAGAGAA AAATAGCAAC	420
	TGATGTGGGT GETTTTTTT TITTTTTAAT TIGAATAAAA AGAATTAGAA GTGATGTCCT	410
40	TTTATAAAAT GECTTCTCCC CCTTCCCGCC TACAGTCTCT TCCTCTCCCC TTAGAGGGGG	540
40	GAAAGTGTAT AAACCTACAG GGTTGTGAGT CTGAAAAGAG GATCCCCCTC ACCCCCACCC	600
	TGGCCAGAGC ASTREGGTT GERGGGTGGG AGAGGGGGAC ACAGATCCTG GCACACTGTG	6 € C
45	GATATTTCTT GCAGATTGCA GTCTCTTGTG GCCCAAACAG GTTAGGTAGA CTATCGCCTC	710
	TGGCAGGTGC CACCTTTTGG TACCAACATG TTCTGAGGTG TTAGGATTTG GGTTGGGTTT	780
50	TITTIGIPIG ITPITITIT CONTITOGIC TITTITITT TOYCCTIKTA AAGAAAAGCO	840
50	AAAGGCCGCT GTGAGTCCTG GTGGCAGGCT CTCCATGGAT GTAGCATATC GAAGATAATT	900
	TTTATACTGC ATTTTTATGG ATTATTTTGT AATGTGTGAT TCCGTCTGCT GAGGAGGTG	960
55	GAGGGGCTCC AGGGAAAGCC ACCCACCTTC AGTGAGGTTG CTCCCCAGCT GAGCGCACCG	1020
	GGCATGGGAT GTGGAGGCTG GCGACACACC CTGTGCCTCT CCAAGGCTGG GCGCGTGGGG	1080
60	CGTCCAGAGT CTTTCTGGGT CTCAGATGTC CATCTGCCAC CTCTTGTTAA GGCTCTAGCT	1140
UU		

	AGAAGGAGA CTGAGGGTAG AAGAAGTTA TTCCCGAAGA AAAAAAGAAT GAAAAACTCAT	1200
	TGTACTGAAC TGTTTTTATA TTTTTTAAAAG TTACTATTTA AAGCGGACGT CGTGGGTCGA	1260
5	CCCGGGAATT CCCGGACCGG TACTGTCAGG TCTAAC	129€
10	(2) INFORMATION FOR SEQ ID NO: 127:	
15	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 737 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear 	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 127:	
20	GGCANAGTGG AGGCAATGCC AGCTCCAGGA CAGAGGCTCA GGTGCCCAAC GGGCAAGGCA	60
	GCCCAGGGGG CTGTGTCTGT TCAAGTCAGG CTTCCCCGGC CCYTCGCGCA NCAGCGCTTC	120
25	CACGGGCAGC CCGGGGCCCC ACCCCACGCA CTGAAGAGGC CGCCTGGGCT GCCATGGCCC	180
23	TGACCTTCCT GTTGGTGCTG CTCACCCTGG CCACGCTCTG CACACGGCTG CACAGAAACT	240
	TCCGACGCGG GBAGAGCATC TACTGGGGGC CCACAGCGGA CAGCCAGGAC ACAGTGGCTG	300
30	CTGTGCTGAA GCGGAGGCTG CTGCAGCCCT CGCGCGGGGT CAAGCGCTCG CGCCGGAGAC	360
	CCYTCYTCCC GCCCACGCCG GACAGCGGCC CGGAAGGCGA GAGCTCGGAG TGACGGCCTG	420
25	GGACCTGCCA CTGTGGCGTG CGGTCTCCCC GCGCGCGAG GCCGCGAMCT NTGCCACGTC	480
35	GACCGCGCGC NGGCCGCTMC CCTGGTGGCG ATGGCGCGGC ACTGGCGAGC ACTGCGKGGG	540
	CTTTCCTCCT TGTTGGTTGC TGAGTGGGCG GCCAAGGGGA GAAAAGGAGC CGCTTYTGCC	600
40	TCCCTTGCCA AAACTCCGTT TCTAATTAAA TTATTTTTAG TAGAAAAAAA AAAAAAAAA	660
	AAAAAAAAA AAAAAAAAAA TCGAGGGGG GCCCGGTACC CAATTNGCCA	720
45	AATAGCGATC GTATNAA	737
50	(2) INFORMATION FOR SEQ ID NO: 128: (i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 1925 base pairs (B) TYPE: nucleic acid	
55	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 128:	
60	CCCCGCCTCC AAAGCTAACC CTCGGGCTTG AGGGGAAGAR GCTGACTGTA CGTTCCTTCI	60

	ACTOTGECAC	CACTCTCCAG	GCTGCCATGG	GGCCCAGCAC	COCTOTCCTC	ATCTTGTTCC	120
	TTTTGTCATG	GTOGGGACCC	CTCCAAGGAC	AGCAGCACCA	CCTTGTGGAG	TACATGGAAC	180
5	GCCGACTAGO	TGCTTTAGAG	GAACGGCTGG	COCAGTGCCA	GGACCAGAGT	AGTCGGCATG	240
	CTGCTGAGCT	GOGGGACTTO	AAGAACAAGA	TOADDETTER	GCTGBAGGTG	GCAGAGAAGG	300
10	DEEMEEDDDA	ACTCAGAACT	GAGGCCGACA	CCATCTCCGG	GAGAGTGGAT	CGTCTGGAGT	360
10	GGGAGGTAGA	CTATCTGGAG	ACCCAGAACC	CAGCTCTGCC	CTGTGTAGAG	TTTGATGAGA	420
	AGGTGACTGG	AGGCCTGGG	ACCAAAGGCA	AGGGAAGAAG	GAATGAGAAG	TACGATATGG	480
15	TGACAGACTG	TGGCTACACA	ATCTCTCAAG	TGAGATCAAT	GAAGATTCTG	AAGCGATTTG	540
	GTGGCCCAGC	TGGTCTATGG	ACCAAGGATC	CACTGGGGCA	aacagagaag	ATCTACGTGT	€00
20	TAGATGGGAC	ACAGAATGAC	ACAGCCTTTG	TETTECCAAG	GCTGCGTGAC	TTCACCCTTG	660
20	CCATGGCTGC	COGGAAAGCT	TCCCGAGTCC	GEGTECCCTT	CCCCTGGGTA	GGCACAGGGC	720
	AGCTGGTATA	TGGTGGCTTT	CTTTATTTTTG	CTCGGAGGCC	TICTGGAAGA	CCTGGTGGAG	780
25	GTGGTGAGAT	GGAGAACACT	TTGCAGCTAA	TUAAATTCCA	CCTGGCAAAC	CGAACAGTGG	840
	TGGACAGITC	AGTATTCCCA	GCAGAGGGGC	TGATCCCCCC	CTACGGCTTG	ACAGCAGACA	900
30	CCTACATCGA	CCTGGCAGCT	GATGAGGAAG	GTCTTTGGGC	TGTCTATGCC	ACCCGGGAG3	960
50	ATGACAGECA	CTTGTGTCTG	GCCAAGTTAG	ATCCACAGAC	ACTGBACACA	GAGCAGCAGT	1020
	GGGACACACC	ATGTCCCAGA	GAGAATGCTG	AGGCTGCCTT	TKTCATCTGT	GGGACCCTCT	1080
35	ATGTCGTCTA	TAACACCCGT	CCTGCCAGTC	GGGCCCGCAT	CCAGTGCTCC	TTTGATGCCA	1140
	GCGGACCCTG	ACCCCTGAAC	GGGCAGCACT	CCCTTATTTT	CCCCGCAGAT	ATGGTGCCCA	1200
40	TGCCAGCCTC	CGCTATAACC	CCCGAGAACG	CCAGCTCTAT	GCCTGGGATG	ATGGCTACCA	1260
	GATTGTCTAT	AAGCTGGAGA	TGAGGAAGAA	AGAGGAGGAG	GTTTGAGGAG	CTAGCCTTGT	1320
	TTTTTGCATC	TTTCTCACTC	CCATACATTT	OTATATTATA	CCCACTAAAT	TTCTTGTTCC	1380
45	TCATTCTTCA	AATGTGGGCC	AGTTGTGGCT	CAAATCCTCT	ATATTTTTAG	CCAATGGCAA	1440
	TCAAATTCTT	TCAGCTCCTT	TGTTTCATAC	GGAACTCCAG	ATCCTGAGTA	ATCCTTTTAG	1500
50	AGCCCGAAGA	GTCAAAACCC	TCAATGTTCC	CTCCTGCTCT	CCTGCCCCAT	GTCAACAAAT	1560
	TTCAGGCTAA	GGATGCCCCA	GACCCAGGGC	TCTAACCTTG	TATGCGGGCA	GGCCCAGGGA	1620
	GCAGGCAGCA	GTGTTCTTCC	CCTCAGAGTG	ACTTGGGGAG	GGAGAAATAG	GAGGAGACGT	1680
55	CCAGCTCTGT	CCTCTCTTCC	TCACTCCTCC	CTTCAGTGTC	CTGAGGAACA	GGACTTTCTC	1740
	CACATTGTTT	TGTATTGCAA	CATTTTGCAT	TAAAAGGAAA	ATCCAMAAAA	AAAAAAAA	1800
60	АААААА	АААААААА	AAAAAAAA	AAAAAAAA	AAAAAAA	AAAAAAAA	1860

WO 98/39446 PCT/US98/04482

273

	ACTGCG300G CTGTCCCTTC TGTCGTCTTC TCGCA300GT ACCCTTCTGT CGTCTTCTCG	1920
	CAGCC	1925
_	CAOCC	
5		
	(2) INFORMATION FOR SEQ ID NO: 129:	
10	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2713 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear 	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 129:	
	TECTACETTE CCAACCETET GGCATECECA GCACTGATGG TECTGGCATE CAEGGETGAG	6C
20	GCCAGCCGTG ACTGCTTCCA TCCCTTGTCA GCAGCCACGA CCCTTTGGTG TACCTGTYTC	120
	AGTTGACAAG GACGTGCATA TTCCTTTCAC CAACGGTTCC TATACCTTTG CCTCTATGTA	180
	CCATCGGCAA GGTGGGGTGC CAGGCACTTT TGCCAATCGT GATTTCCCCC CTTCTCTACT	240
25	ACACCTCCAC CCTCAATTTG CTCCCCCAAA TCTAGATTGC ACCCCAATCA GTATGCTGAA	300
	TCATAAGTGG TGTGGGGGTT TCCGGCCTTT GECTCCACCC GRGGACCGGG RGAGYTATCA	360
30	GTCAGITTTA CGCCGGCCAA GCGACTTAAG AACTBICATG ACACAGAGTC TCCCCACTTG	420
	CGCNTCTCAG ATGCAGATGG GAANGAATAT GACTTTGGGA CACAGCTGCM ATCTAGCTCC	480
	CCCGGTTCAC TAAAGGTTGA TGACACTGGG AAGAAGATTT TTGCTGTCTC TGGCCTCATT	5 4 0
35	TCTGATCGGG AAGCCTCATC TAGCCCAGAG GNTCGGNAAT GACAGATGTA AGAAGAAAGC	600
	AGCGGCATTG TTCGACAGCC AGGCCCCAAT TYGCCCCATC TGCCAGGTCC TGCTGAGGCC	660
40	CAGTGAGCTG CAGGAGCATA TGGAGCAGGA ACTGGAGCAG CTAGCCCAAC TGCCCTCGAG	720
	CAAGAATTCC CTTCTGAAGG ATGCCATGGC TCCAGGCACC CCAAAGTCCC TCCTGTTGTC	780
	TGCTTCCATC AAGAGGGAAG GAGAGTCTCC AACGGCATCA CCCCACTCAT CTGCCACCGA	840
45	TGACCTCCAC CATTCAGACA GATACCAGAC CTTTCTGCGA GTACGAGCCA ACCGGCAGAC	900
	CCGAYTGAAT GYTCGGATTG GGAAAATGAA ACGGAGGAAG CAAGATGAAG GGCAGGTATG	960
50	TCCCCTGTGC AACCGCCCCC TGGCAGGATC GBAGCAGGAG ATGAGTAGGC ATGTGGAGCA	1620
23	TTGCCTTTCT AAGAGGGAAG GCTCCTGCAT GGCTGAGGAT GATGCTGTGG ACATCGAGCA	10:80
	TGAGAACAAC AACCGCTTTG AGGAGTATGA GTGTTGTGGA CAGAAGCGGA TACGGGCCAC	1140
55	CACTOTOCTG GAAGGTGGCT TOOGAGGCTC TOGGTTCATC ATGTGCAGCG GCAAAGAGAA	1200
	CCCGGACAGT GATGCTGACT TGGATGTGGA TGGGGATGAC ACTCTGGAGT ATGGGAAGCC	1260
60	ACAATACACA GAGGCTGATG TCATCCCCTG CACAGGCGAG GAGCCTGGTG AAGCCAAGGA	1320
00	A A COM MI A A A PORTO TO THE TOTAL TOTAL TO THE TOTAL TOTAL TO THE TO	

Control of the Contro

	GAGAGAGGCA	. CTTCGGGGGG	CAGICCTAAA	TGGCGCCCT	CCCAGCACGC	GCATCACACC	1380
5	TGAGTTCTCT	AAATGGGCCA	GTGATGAGAT	GCCATCCACC	AGCAATGGTG	AAAGCAGCAA	1440
·	GCAGGAGGCC	ATGCAGAAGA	CCTGCAAGAA	CAGCGACATC	GAGAAAATCA	CCGAAGATTC	1500
	AGCTGTGACC	ACGTTTGAGG	CTCTGAAGGC	TCGGGTCAGA	GAACTTGAAC	GGCAGCTATC	1560
10	TOGTGGGGAC	CGTTACAAAT	GCCTCATCTG	CATGGACTCG	TACTCGATGC	CCCTAACGTC	1620
	CATCCAGTGT	TGGCACGTGC	ACI GCGAGGA	GTGCTGGCTG	CGGACCCTGG	GTGCCAAGAA	1680
15	GCTCTGCCCT	CAGTGCAACA	CGATCACAGC	GCCCGGAGAC	CTGCGGAGGA	TCTACTTGTG	1740
15	AGCTATCTGC	CCCAGGCAGG	CCTCGCCTCC	AGCAGCCCCA	CCTGCCCCCA	GCCTCTGTGA	1800
	CAGTGACCGT	YTCCCTTTGT	ACATACTIGC	ACACAGGTTC	CCCATGTACA	TACATGCACA	1860
20	TACTCAAACA	TGCGTACACA	CACACACATT	TACACACGCA	GGACTCTGGA	GCCAGAGTAG	1920
	AGGCTGTGGC	CCAGGCACTA	CCTGCTGGCT	CCCACCTATG	GTTTGGGGGC	CATACCTGTT	1980
25	CCAGCTCTGT	TCCCAGGGTG	GGGCAGGGAG	GTGGGGGTTG	GGGGAGTAGT	GGGGCACGGC	2040
	TCCTAAGATC	CAGCCCCCAT	ACTGACAGAC	GGACAGACAG	ACATGCAAAC	ACCAGACTGA	2100
	AGCACATGTA	ATATAGACCG	TGTATGTTTA	CAATGTTGTG	TATAAATGGG	ACAACTCCTC	2160
30	GCCCTCTACC	TGTCCCCTCC	CCCTTTGGTT	GTATGATTTT	CTTCTTTTTT	AAGAACCCCT	2220
	GGAAGCAGCG	CCTCCTTCAG	GGTTGGCTGG	GAGCTCGGCC	CATCCACCTC	TTGGGGTAYC	2280
35	TGCCTCTCTC	TCTCCTGTGG	TGTCCCTTCC	CTCTCCCATG	TGCTCGGTGT	TCAGTGGTGT	1340
	ATATTTCTTC	TCCCAGACAT	GGFFCACACG	CCCCAAGGGA	CATGATCCTC	TCCTTAGTCT	2400
	TAGCTCATGG	GGCTCTTTAT	AAGGAGTTGG	GGGGTAGAGG	CAGGAAATGG	GAACCGAGCT	2.460
40	GAAGCAGAGG	CTGAGTTAGG	ODADATCEOD	ACAGTGCTCC	TGGCCACCCA	GCCTCTGCTG	2520
	AGAACCATTC	CTGGGATTAG	AGCTGCCTTT	CCCAGGGAAA	AAGTGTCGTC	TCCCCGACCC	2580
45	TCCCGTGGGC	CCTGTGGTGT	GATGCTGTGT	CTGTATATTC	TATACAAAGG	TACTTGTCCT	2640
	TTCCCTTTGT	AAACTACATT	TGACATGGAT	TAAACCAGTA	TAAACAGTTA	AAAAAAAA	2700
	AAAAAAAACT	CGA					2713
50							

(2) INFORMATION FOR SEQ ID NO: 130:

55 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1011 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

PCT/US98/04482

The Control of Security A.

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 130:	
	AGAGGAGGT GTGACCCGG AGGAAGTAGA GCCTGAGGAG GCATCTCTCTGA	€0
5	GCAACCCTGC CCAGCTGACA CAGAGGTGGT GGAAGACTCC TTGAGGCAGC GTAAAAGTCA	120
	GCATGCTGAC AAGGAACTGT AGATTTAATG ATTGCGTTTTC AAGAATACAC ACCAAAACAA	180
10	TATGTCAGCT TCCCTTTGGC CIGCAGTTTG TACCAAATCC TTAATTTTTY YTGAATGAGC	240
10	AAGCTTCTCT TAAAAGATGC TCTCTAGTCA TTTGGTCTCA TGGCAGTAAG CCTCATGTAT	300
	ACTAAGGAGA GTCTTCCAGG TGTGACAATC AGGATATAGA AAAACAAACG TAGTGTNTGG	360
15	GATCTGTTTG GAGACTGGGA TGGGAACAAG TTCATTTACT TAGGGGTCAG AGAGTCTCGA	420
	CCAGAGGAGG CCATTCCCAG TCCTAATCAG CACCTTCCAG AGACAAGGCT GCAGGCCCTG	480
	TGAAATGAAA GCCAAGCAGG AGCCTTGGCT CTGAGNCATC CCCAAAGTGT AACGTAGAAG	540
20	OPPROVIDE TOTAL ATTITUTE AND ATTOCAMENT OF THE TOTAL ACCURACY.	600
	ACCACAGTGC ATGAAAAATC TTTCACAGCT AGAAATTGAA AGGGCCTTGG GTATAGAGAG	660
25	CAGCTCAGAA GTCATCCCAG CCCTCTGAAT CTCCTGTGCT ATGTTTTATT TCTTACCTTT	720
	AATTITTCCA GCATTTCCAC CATGGGCATT CAGGCTCTCC ACACTCTTCA CTATTATCTC	780
• •	TIGGTCAGAG GACTCCAATA ACAGCCAGGT TTACATGAAC TGTGTTTGTT CATTCTGACC	840
30	TAAGGGGTTT AGATAATCAG TAACCATAAC CCCTGAAGCT GTGACTGCCA AACATCTCAA	900
	ATGAAATGTT GTRGCCATCA GAGACTCAAA AGGAAGTAAG GATTTTACAA GACAGATTAA	960
35	AAAAAATTG TTTTGTCCAA AAAANAAAAA AAAAAAAACTC GAAGGGGGGG C	1011
40		
40	(2) INFORMATION FOR SEQ ID NO: 131:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2278 base pairs	
45	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 131:	60
50	GTAATTCGGC ACGAGGCGCC CAACATGGCG GGTGGGCGCT GCGGCCCGCA SCTAACGGCG	120
	CTCCTGGCCG CCTGGATCGC GGCTGTGGCG GCGACGCAG GCCCCGAGGA GGCCGCGCTG	180
55	CCGCCGGAGC AGAGCCGGGT CCAGCCCATG ACCGCCTCCA ACTGGACGCT GGTGATGGAG	240
	GGCGAGTGGA TGCTGAAATT TTACGCCCCA TGGTGTCCAT CCTGCCAGCA GACTGATTCA	
.	GAATGGGAGG CTTTTGCAAA GAATGGTGAA ATACTTCAGA TCAGTGTGGG GAAGGTAGAT	300 360
60	GTCATTCAAG AACCAGGTTT GAGTGGCCGC TTCTTTGTCA CCACTCTCCC AGCATTTTTT	200

	CATGCAAAGG	ATGGGATATT	CCGCCGTTAT	CGTGGCCCAG	GAATOTTOGA	AGACCTGCAG	420
5	AATIATATCT	TAGAGAAGAA	ATGGCAATCA	STOGAGOSTS	TGACTGGCTG	GAAATCCCCG	480
ن	GCTTCTCTAA	CGATGTCTGG	AATGGCTGGT	CTTTTTAGCA	TOTOTGGGGAA	GATATGGCAT	540
	CTICACAACT	ATTTCACACT	GACTCTTGGA	ATTOCTGOTT	GGTGTTCTTA	TGTCTTTTTC	600
10	GTCATAGCCA	CCTTGGTTTT	TYGGOOTTTTT	ESTOTESSTA	TOTTSTTGGT	AATATCAGAA	660
	TGTTTCTATG	TGCCACTTCC	AAGGCATTTA	TOTGAGOGTT	CTGAGCAGAA	TOGGAGATCA	720
15	GAGGAGGCTC	ATAGAGCTGA	ACAGTTGCAG	GATGCGGAGG	AGGAAAAAGA	TGATTCAAAT	780
	GAAGAAGAAA	ACAAAGACAG	CCTTSTAGAT	GATGAAGAAG	AGAAAGAAGA	TOTTGGCGAT	840
	GAGGATGAAG	CAGAGGAAGA	AGAGGAGGAG	GACAACTTGG	CTSCTGGTGT	GGATGAGGAG	900
20	AGAAGTGAGG	CCAATGATCA	GGGGCCCA	GGAGAGGACG	GTGTGACCCG	GGAGGNAAGT	960
	AGAGCCTGAG	GAGGTTGAAG	OTOTACEDAA	TGAGCAACCC	TGCCCAGCTG	ACACAGAGGT	1020
25	GGTGGAAGAC	TCCTTGAGGC	AGIGTAAAAG	TCAGCATGCT	GNCAAGGGAC	TGTAGATTTA	1080
	ATGATGCGTT	TTCAAGAATA	CACACCAAAA	CAATATGTCA	GCTTCCCTTT	GGCCTGCAGT	1140
	TTGTACCAAA	TOCTTAATTT	TTCCTGAATG	AGCAAGCTTC	TOTTAAAAGA	TGCTCTCTAG	120
30	TCATTTGGTC	TCATGGCAGT	AAGCCTCATG	TATACTAAGG	AGAGTCTTCC	AGGTGTGACA	1260
	ATCAGGATAT	AGAAAACAA	ACGTAGTGTN	TGGGATCTGT	TTGGAGACTG	GGATGGGAAC	1320
35	AAGTTCATTT	ACTTAGGGGT	CAGAGAGTCT	CGACCAGAGG	AGGCCATTCC	CAGTCCTAAT	1380
	CAGCACCTTC	CAGAGACAAG	GCTGCAGGCC	ADTAAAETET	AAGCCAAGCA	GGAGCCTT G G	1440
	CTCTGAGGCA	TCCCCAAAGT	GTAACGTAGA	AGCCTTGCAT	COTTTTCTTG	TGTAAAGTAT	15(0)
40	TTATTTTTTGT	CAAATTGCAG	GAAACATCAG	GCACCACAGT	GCATGAAAAA	TCTTTCACAG	1560
	CTAGAAATTG	AAAGGGCCTT	GGGTATAGAG	AGCAGCTCAG	AAGTCATCCC	AGCCCTCTGA	1620
45	ATCTCCTGTG	CTATGTTTA	TTTCTTACCT	TTAATTTTTC	CAGCATTTCC	ACCATGGGCA	1680
	TTCAGGCTCT	CCACACTCTT	CACTATTATC	TCTTGGTCAG	AGGACTCCAA	TAACAGCCAG	1740
	GTTTACATGA	ACTGTGTTTG	TICATTCTGA	COTAAGGGGT	TYPAGATAATC	AGTAACCATA	1800
50	ACCOCTGAAG	CTGTGACTGC	CAAACATCTC	AAATGAAATG	TESTRECCAT	CAGAGACTCA	1860
	AAAGGAAGTA	AGGATTTTAC	AAGACAGATT	TAAAAAAA	TETTTTGTCC	NAAAATATAG	1920
55	TTGTTGTTGA	TTTTTTTTTA	AGTTTTCTAA	GCAATATTTT	TCAAGCCAGA	AGTCCTCTAA	1980
	GTCTTGCCAG	TACAAGGTAG	TOTTGTGAAG	AAAGTTGAA	TACTGTTTTG	TTTTCATCTC	2040
	AAGGGGTTCC	CTGGGTCTTG	AACTACTTTA	ATAATAACTA	AAAAACCACT	TCTGATTTTC	2100
60	CTTCAGTGAT	GTGCTTTTGG	TGAAAGAATT	AATGAACTCC	AGTACCTGAA	AGTGAAAGAT	2160

	TTGATTTTGT	TTCCATCTTC	TGTAATCTTC	CAAAGAATTA	TATCTTTGTA	AATCTCTCAA	2220
5	TACTCAATCT	ACTGIAAGTA	CCCAGGGRGG	STAATTTCYT	AAAAAAT	ААААААА	2278
10		STION FOR SE		ICS:			
15		(B) TYP (C) STR	E: nucleic ANDEDNESS: OLOGY: line	acid double			
	(xi)) SEQUENCE I	DESCRIPTION	: SEQ ID NO	: 132:		
20	GGCAGGGGCG	GCGTGAACCC	GTCGGGCACT	GTGTCCCTGA	CAATGGGAAC	AGCCGACAGT	60
20	GATGAGATGG	CCCCGGAGCC	CCACAGCACA	CCCACATCGA	TGTGCACATC	CACCAGGAGT	120
	CTGCCCTGGC	CAAGCTCCTG	CTCACCTGCT	GCTCTGCGCT	GCGGCCCCGG	GCCACCCAGG	180
25	CCAGGGGCAG	CANCCGGCTG	CTGGTGGCCT	CGTGGGTGAT	GCAGATCGTG	CTGGGGATCT	240
	TGAGTGCAGT	CCTAGGAGGA	TTTTTCTACA	TCCGCGACTA	CACCCTCCTC	GTCACCTCGG	300
20	GAGCTGCCAT	CTGGACAGGG	GCTGTGGCTG	TGCTGGCTGG	AGCTGCTGCC	TTCATTTAYG	360
30	AGAAACGGGG	TGGTACATAC	TGGGCCCTGC	TGAGGACTCT	GCTARCGCTG	GCAGCTTTCT	420
	CCACAGCCAT	CGCTGCCCTC	AAACTTTGGA	ATGAAGATTT	CCGATATGGC	TACTCTTATT	480
35	ACAACAGTGC	CTGCCGCATC	TCCAGCTCGA	GTGACTGGAA	CACTCCAGCC	CCCACTCAGA	540
	GTCCAGAAGA	AGTCAGAAGG	CTACACCTAT	GTACCTCCTT	CATGGACATG	CTGAAGGCCT	600
	TGTTCAGAAC	CCTTCAGGCC	ATGCTCTTGG	GTGTCTGGAT	TCTGCTGCTT	CTGGCATCTC	660
40	TGGCCCCTCT	GTGGCTGTAC	TGCTGGAGAA	TGTTCCCAAC	CAAAGGGAAA	AGAGACCAGA	720
	AGGAAATGTT	GGAAGTGAGT	GGAATCTAGC	CATGCCTCTC	CTGATTATTA	GTGCCTGGTG	780
45	CTTCTGCACC	GGGCGTCCCT	GCATCTGACT	GCTGGAAGAA	GAACCAGACT	GAGGAAAAGA	840
	GGCTCTTCAA	CAGCCCCAGT	TATCCTGGCC	CCATGACCGT	GGCCACAGCC	CTGCTCCAGC	900
						CCCCTCCTGA	960
50						TAAAAAAA	
<i>5 F</i>		CCGG-ACCCA	1 TGGGCCTNN	GOGGGINGG I.T.	, AAAATTAAT	GGGGGGGTT	
55	TAAAAGGG						1088

Alternative Contract Contract Andrews

⁶⁰ (2) INFORMATION FOR SEQ ID NO: 133:

5	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 553 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: doubl€ (D) TOPOLOGY: linear 	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 133:	
10	GGCAGAGAGC AGATGGCCTT GACACCAGCA GGCTGACATC CGCTATTGCT ACTTCTCTG	60
	TCCCCCACAG TTCCTCTGGA CTTCTCTGGA CCACAGTCCT CTGCCAGACC CCTGCCAGAC	120
1.5	CCCAGTCCAC CATGATCCAT CTGGGTCACA TCCTCTTCCT GCTTTTGCTC CCAGTGGCTC	180
15	CAGCTCAGAC GACTCCAGGA GAGAGATCAT CACTCCTGC CTTTTACCCT GGCACTTCAG	240
	GCTCTTGTTC CGGATGTGGG TCCCTCTCTC TGCCGCTCCT GGCAGGCCTC GTGGCTGCTG	300
20	ADSCRIPT DESCRIPTION OF ADSCRIPTION	360
	GCCCCGCCCA AGATGGCAAA GTCTACATCA ACATGCCAGG CAGGGGCTGA CCCTCCTGCA	420
25	GCTTGGACCT TTGACTTCTG ACCCTCTCAT CCTGGATGGT GTGTGGTGGC ACAGGAACCC	480
23	CCGCCCCAAC TTTIGGATTG TAATAAAACA ATTGAAACAC CAAAAAAAAA AAAAAAAAAA	540
	AAA AAAAAAA	553
30		
	(2) INFORMATION FOR SEQ ID NO: 134:	
35	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 467 amino acids (B) TYPE: amino acid	
	(D) TOPOLOGY: linea: (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 134:	
40	Met Arg Pro Gln Glu Leu Pro Arg Leu Ala Phe Pro Leu Leu Leu Leu 15 15	
45	Leu Leu Leu Leu Pro Pro Fro Pro Cys Pro Ala His Ser Ala Tha 20 25 30	
	Arg Phe Asp Pro Thr Trp Glu Ser Leu Asp Ala Arg Gln Leu Pro Ala 35 40 45	
50	Trp Phe Asp Gln Ala Lys Phe Gly Ile Phe Ile His Trp Gly Val Phe 50 60	
55	Ser Val Pro Ser Phe Gly Ser Glu Trp Phe Trp Trp Tyr Trp Gln Lys 65 70 75 80	
JJ	Glu Lys Ile Pro Lys Tyr Val Glu Phe Met Lys Asp Asn Tyr Pro Pro 85 90 95	
60	Xaa Phe Lys Tyr Glu Asp Phe Gly Pro Leu Phe Thr Ala Lys Phe Phe	

Les Company (Company Laber)

	Asr.	Ala	Asn 115	Gln	Trp	Ala	Xaa	11e 120	Pn€	Gin	Ala	Ser	125	Ala	Lys	ТУІ
5	Ile	Val 130	Leu	Thr	Ser	Lys	His 135	His	Glu	Gly	Phe	Thr 140	Leu	Trp	Gly	Se:
10	Glu 145	Tyr	Ser	Trp	Asn	Trp 150	Asn	Ala	Ile	Asp	Glu 155	Gly	Pro	Lys	Arg	Asr 160
10	lie	Val	Lys	Glu	Leu 165	Glu	Val	Ala	Il ϵ	Arg 170	Asn	Arg	Thr	Asp	Leu 175	Arg
15	Phe	Gly	Leu	Tyr 180	Tyr	Ser	Leu	Phe	Glu 185	Trp	Phe	His	Pro	Leu 190	Phe	Leu
	Glu	Asp	Glu 195	Ser	Ser	Ser	Phe	His 200	Lys	Arg	Gln	Phe	Pro 205	Val	Ser	Lys
20	Thr	Leu 21(Pro	Glu	Leu	Tyr	Glu 215	Leu	Val	Asn	Asn	Tyr 220	Gln	Pro	Glu	Val
25	Leu 225	Trp	Ser	Asp	Gly	Asp 230	Gly	Gly	Ala	Pro	Asp 235	Gln	Tyr	Trp	Asn	Xaa 240
23	Thr	Gly	Phe	Leu	Ala 245	Trp	Leu	Tyr	Asn	Glu 250	Ser	Pro	Val	Arg	Gly 255	The
30	Val	Val	Thr	Asn 260	Asp	Arg	Trp	Gly	Ala 265	Gly	Ser	Ile	Cys	Lys 270	His	Gly.
	Gly	Phe	Tyr 275	Thr	Cys	Ser	Asp	Arg 280	Tyr	Asn	Pro	Gly	His 285	Leu	Leu	Pro
35	His	Lys 290		Glu	Asn	Cys	Met 295		Ile	Asp	Lys	Leu 300	Ser	Trp	Gly	Tyr
40	Arg 305		Glu	Ala	Gly	Ile 310		Asp	Tyr	Leu	Thr 315		Glu	Glu	Leu	Val 320
40	Lys	Gln	Leu	Val	Glu 325	Thr	Val	Ser	Cys	Gly 33(Asn	Leu	Leu	Met 335	
45	lle	Gly	Prc	Thr 340		Asp	Gly	Thr	11e 345		· Val	. Val	Phe	Glu 350		Arg
	Leu	Arg	Gln 355		Gly	Ser	Trp	360		Val	. Asr	Gly	Glu 365		. Ile	Tyr
50	Glu	370		Thr	Trp	Arg	Ser 375		. Asn	Asp	Thi	7 Val	Thr	Pro	Asp	Val
55	Trp 385		Thr	Ser	Lys	Pro 390		s Glu	Lys	: Lev	399		Ala	ılle	Ph€	1€u 400
JJ	Lys	Trp	o Pro	Thi	Ser 405	Gly	/ Glr	n Leu	ı Ph∈	Let 410		y His	s Pro	Lys	Ala 415	
60	Lev	ı Gly	y Ala	a Thi 420		ı Val	Ly	s Leu	Let 425		y Hi	s Gly	y Glr	430		a Asr.

	Trp	p Ile	e Sei 435	: Lev	ı Glu	ı Glr	ı Asr	n Gly 44(e Met	Val	Glu	1 Leu 445		Glr.	Leu
5	Tha	r Ile 450	∈ His	Glr	i Met	: Pro	Cys 455		Trp	Gly	, Trp	Ala 460		Ala	Leu	Thr
10	Asr 465		l Il∈	-												
	(2)	INF	FORMA	TION	FOR	. SEQ) ID	N O:	135:							
15					(A) I (B) 7 (D) 7	LENGT TYPE : TOPOI	TH: 2 : am: LOGY:	TERIS 222 a inc a : lir	mind cic mear	aci		: 13	5:			
20	Met I	Trp	Ser	Ala	Gly	Arg	Gly	Gly	Ala	Ala 10		Pro	Val	Leu	Leu 15	Gly
25	Leu	Leu	. Leu	Ala 20		Leu	Val	Pro	Gly 25	Gly	Gly	Ala	Ala	Lys 30	Thr	Gly
	Ala	Glu	Leu 35	Val	Thr	Сув	Gly	Ser 4(Val	Leu	Lys	Leu	Leu 45	Asn	Thr	His
30	His	Arg 50	Val	Arg	Leu	His	Ser 51	His	Asp	Ile	Lys	Tyr 60	Gly	Ser	Gly	Ser
35	Gly 65	Gln	Gln	Ser	Val	Thr 70	Gly	Vā]	Glu	Ala	Ser 75	Asp	Asp	Ala	Asn	Ser 80
	Tyr	Trp	Arg	Ile	Arg 85	Gly	Gly	Ser	Glu	Gly 90	Gly	Cys	Arg	Arg	Gly 95	Sei
40	Fro	Val	Arg	Cys 100	Gly	Gln	Ala	Val	Arg 105	Leu	Thr	His	Val	Leu 110	Thr	Gly
	Lys	Asn	Leu 115	His	Thr	His	His	Phe 120	Pro	Ser	Pro	Leu	Ser 125	Asn	Asn	Gln
45	Glu	Val 130	Ser	Ala	Phe	Gly	Glu 135	Asp	Gly	Glu	Gly	Asp 140	Asp	Leu	Asp	Leu
50	Trp 145	Thr	Val	Arg	Cys	Ser 150	Gly	Gln	His	dıņ	Glu 155	Arg	Glu	Ala	Ala	Val 160
	Arg	Phe	Gln	His	Val 165	Gly	Thr	Ser	Val	Phe 170	Leu	Ser	Val	Thr	Gly 175	Glu
55	Gln	Tyr	Gly	Ser 180	Pro	Ile	Arg	Gly	Gln 185	His	Glu	Val	His	Gly 190	Met	Prc
	Ser	Ala	Asn 195	Thr	His	Asn	Thr	Trp 200	Lys	Ala	Met	Glu	Gly 205	Ile	Phe	Ile
60	Lys	Pro	Ser	Val	Glu	Pro	Ser	Ala	Gly	His	Asp	Glu	Leu	Xaa		

WO 98/39446 PCT/US98/04482

215

210

60

281

226

5 (2) INFORMATION FOR SEQ ID NO: 136: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 156 amino acids (B) TYPE: amino acid 10 (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 136: Met Val Ile Glu Ile Ser Asn Lys Thr Ser Ser Ser Thr Cys Ile 5 10 15 Leu Val Leu Leu Val Ser Phe Cys Leu Leu Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro Ala Glu His Gly Val Leu Ser 20 Arg Gln Leu Arg Ala Leu Pro Ser Glu Asp Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys Asp Ser Thr His Gln Trp Leu 25 Asp Gly Ser Asp Cys Val Leu Gln Ala Pro Gly Asn Thr Ser Cys Leu 90 30 Leu His Tyr Met Pro Gln Ala Pro Ser Ala Glu Pro Pro Leu Glu Trp 105 100 Pro Phe Pro Asp Leu Phe Ser Glu Pro Leu Cys Arg Gly Pro Ile Leu 35 120 Pro Leu Gln Ala Asn Leu Thr Arg Lys Gly Gly Trp Leu Pro Thr Gly 40 Ser Pro Ser Val Ile Leu Gln Asp Arg Tyr Ser Gly (2) INFORMATION FOR SEQ ID NO: 137: 45 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 233 amino acids (B) TYPE: amino acid 50 (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 137: Met Met Ile Leu Phe Asn Leu Leu Ile Phe Leu Cys Gly Ala Ala Leu 10 55 Leu Ala Val Gly Ile Trp Val Ser Ile Asp Gly Ala Ser Phe Leu Lys 2.0 Ile Phe Gly Pro Leu Ser Ser Ser Ala Met Gln Phe Val Asn Val Gly

	Tyr	Phe 50	Leu	Πe	Ala	Ala	2F GJA	Val	Val	Val	Phe	Ala 60	Leu	Gly	Phe	Leu
5	Gly 65	Cys	Tyr	Gly	Ala	Lys 70	Thr	Glu	Ser	Lys	Cys 75	Ala	Leu	Vā]	Thr	Phe 80
10	Fhe	Phe	Il€	Leu	Leu 85	Leu	Ile	Ph∈	Il€	Ala 90	Glu	Val	Ala	Ala	Ala 95	Val
•	Val	Ala	Leu	Val 100	Тут	Thr	Thr	Met	Ala 105	Glu	His	Phe	Leu	Thr 110	Leu	Leu
15	Val	Val	Pro 115	Ala	Il€	Lys	Lys	Asp 120	Tyr	Gly	Ser	Gln	Glu 121	Asp	Phe	Thr
	Gln	Val 130	Trp	Asn	Thr	Thr	Met 135	Lys	Gly	Leu	Lys	Cys 1 4 0	Cys	Gly	Phe	Thr
20	Asr. 145	Tyr	Thr	Asp	Phe	Glu 150	Asp	Ser	Pro	Tyr	Phe 155	Lys	Glu	Asn	Ser	Ala 160
25	Phe	Pro	Prc	Phe	Суs 165	СЛЕ	Asn	Asp	Asn	Val 170	Thr	Asn	Thr	Ala	Asn 175	Glu
<i>20</i>	Thr	Cys	Thr	Lys 180	Gln	Lys	Ala	His	Asp 185	Gln	Lys	Val	Glu	Gly 19(Cys	Phe
30	Asr.	Gln	Leu 195	Leu	Tyr	Asp	Ile	Arg 200	Thr	Asn	Ala	Val	7hr 201	Val	Gly	Gīу
	Val	Ala 210		Gly	Пe	Gly	Gly 215	Leu	Glu	Leu	Ala	Ala 22(Met	Ile	Val	Ser
35	Met 225	Tyr	Leu	Tyr	Cys	Asn 23(-	Leu	Gln	Xaa							
40	(2)	INF	orma'	TION	FOR	SEQ	ID 1	N O:]	138:							
			(i)	-				ERIS			is					
45			(xi)	(D) I	CPOL	.OGY :	no a lin PTIO	ear	EQ I	D NC	: 13	٤:			
50	Met 1	Gly	Ser	Ser	f Yra	Trp	Ser	Va]	Ala	Cys 10	Pro	Thr	Gly	Leu	Gly 15	Val
50	Leu	Met	Leu	Gly 20	Leu	Gly	Gly	Asp	His	Pro	Pro	Gly	Ser	Gln 30	Val	Asp
55	Pro	Leu	Leu 35		Gly	Хаа	Cys	Val 40	_	Pro	Xaa	Leu	Pro 45	Glu	Leu	Thr
	Ala	Xaa 50	Trp	Arg	Glu	Xaa	Gln 55	Xaa	Arg	Ser	Ala	Ser 60				
60																

WO 98/39446 PCT/US98/04482

283

```
(2) INFORMATION FOR SEQ ID NO: 139:
             (i) SEQUENCE CHARACTERISTICS:
5
                   (A) LENGTH: 73 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 139
     Met Gly Trp Leu Phe Leu Lys Val Leu Leu Ala Gly Val Ser Phe Ser
10
     Gly Phe Leu Tyr Pro Leu Val Asp Phe Cys Ile Ser Gly Lys Thr Arg
15
     Gly Gln Lys Pro Asn Phe Val Ile Ile Leu Ala Asp Asp Met Gly Trp
                                   40
      Gly Asp Trp Gly Ala Asn Trp Ala Glu Thr Lys Asp Thr Ala Asn Leu
20
                              5.5
      Asp Lys Met Ala Ser Glu Gly Met Xaa
                          7 C
      65
25
      (2) INFORMATION FOR SEQ ID NO: 140:
             (i) SEQUENCE CHARACTERISTICS:
30
                    (A) LENGTH: 377 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 140:
      Met His Gly Asn Glu Ala Leu Gly Arg Glu Leu Leu Leu Leu Met
35
      Gln Phe Leu Cys His Glu Phe Leu Arg Gly Asn Pro Arg Val Thr Arg
40
      Leu Leu Ser Glu Met Arg Ile His Leu Leu Pro Ser Met Asn Pro Asp
      Gly Tyr Glu Ile Ala Tyr His Arg Gly Ser Glu Leu Val Gly Trp Ala
45
      Glu Gly Arg Trp Asn Asn Gln Ser Ile Asp Leu Asn His Asn Phe Ala
      Asp Leu Asn Thr Pro Leu Trp Glu Ala Gln Asp Asp Gly Lys Val Pro
50
                                           90
                       85
      His Ile Val Pro Asn His His Leu Pro Leu Pro Thr Tyr Tyr Thr Leu
                                      105
                  100
55
      Pro Asn Ala Thr Val Ala Pro Glu Thr Arg Ala Val Ile Lys Trp Met
      Lys Arg Ile Pro Phe Val Leu Ser Ala Asn Leu His Gly Gly Glu Leu
                            135
60
```

Charles of Lagran Granulars

	Val 145	Val	Ser	Тут	Prc	Phe 150	Asp	Met	Thr	Arg	Thr 15!	Pro	Trp	Alā	Ala	Arç 1€0
5	Glu	Leu	Thr	Frc	Thr 165	Prc	Asp	Asp	Ala	Val 17(Ph€	Arg	Trp	Leu	Ser 175	Thr
10	Val	Tyr	Ala	Gly 18(Ser	Asn	Leu	Alā	Met 185	Gln	Asp	Thr	Ser	Arg 190	Arg	Prc
10	Сує	His	Ser 195	Gln	Asp	Phe	Ser	Val 200	His	Gly	Asn	Ile	lle 201	Asn	Gly	Ala
15	Asp	Trp 210	His	Thr	Val	Pro	Gly 215	Ser	Met	Asn	Asp	Phe 220	Ser	Tyr	Leu	His
	Thr 225	Asn	CAE	Phe	Glu	Val 230	Thr	Val	Glu	Leu	Ser 235	Суѕ	Asp	Lys	Phe	Prc 240
20	His	Glu	Asn	Glu	Leu 245	Pro	Gln	Glu	Trp	Glu 25(Asn	Asn	Lys	Asp	Ala 255	Leu
25	Leu	Thr	Tyr	Leu 260	Glu	Gln	Val	Arg	Met 265	Gly	Ile	Ala	Gly	Val 270	Val	Arg
	Asp	Lys	Asp 275	Thr	Glu	Leu	Gly	11e 280	Ala	Asp	Ala	Val	11e 285	Ala	Val	Asp
30	Gly	11e 290	Asn	His	Asp	Val	Thr 295	Thr	Ala	Trp	Gly	Gly 30(Asp	Tyr	Trp	Arç
	Leu 305	Leu	Thr	Pro	Gly	Asp 310		Met	Val	Thr	Ala 315	Ser	Ala	Glu	Gly	Tyr 32(
35	His	Ser	Val	Thr	Arg 325	Asn	CAE	Arg	Val	Thr 330	Phe	Glu	Glu	Gly	Pro 335	Phe
40	Prc	Сує	Asn	Phe 34(Val	Leu	Thr	Lys	Thr 345	Pro	Lys	Gln	Arg	Leu 350	Arg	Glu
,,	Leu	Leu	Ala 355	Ala	Gly	Ala	Lys	Val 360	Pro	Pro	Asp	Leu	Arg 365	Arg	Arg	Leu
45	Glu	Arg 370		Arg	Gly	Gln	Lys 375	Asp	Xaa							
	(2)	TME	AMA	T ON	FOR	SEC	e TD	NO:	141:							
50	(2)	2111		SEQU	JENCE	CHA	.RAC	TERIS	TICS		ic					
55			/ - - 1		(B) 1	TYPE TOPOI	: am: LOGY	ino a : lir	aciā near			. 1/				
55				·	ıle	. Lev		IPTIC Leu		Ala	Val			e Leu		Ser
60	1 Leu		: Val	Val	His		ı Phe	e Gln	Ile	10 Lev		Leu	ser	Gly	15 Thr	

285

30 Tyr Fro Lys Fhe Tyr Gln Thr Leu His Arg Gln. 5 (2) INFORMATION FOR SEQ ID NO: 142: 10 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 41 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 142: 15 Met Val His Val Leu Glu Ile Leu Leu Phe Ile Thr Met Gln Ala Val Ser Phe Fro Phe Gln Thr Gln Ile Asp Thr Cys Asn Thr Gln Asp Pro 20 25 Ala Glu Arg Gln Pro Ala Ser Ile Val 35 4 C 25 (2) INFORMATION FOR SEQ ID NO: 143: (i) SEQUENCE CHARACTERISTICS: 30 (A) LENGTH: 70 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 143: 35 Met Gly Ser Cys Ser Lys Asn Arg Ser Phe Phe Trp Met Thr Gly Leu Leu Val Phe Ile Ser Leu Leu Ser Glu Trp Gln Gly Pro Trp Glu 25 40 Gly Arg Ala Ile Gly Glu Gly Trp Ala Ser Trp Ala Leu Thr Asn Gly 40 Trp Ala Val Gln Leu Leu Met Ser Leu Gly Asn Asn Thr Glu Lys His 45 Ser Val Met Ile Tyr Glu 50 (2) INFORMATION FOR SEQ ID NO: 144: (i) SEQUENCE CHARACTERISTICS: 55 (A) LENGTH: 483 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 144: 60 Met Ala Thr Gly Gly Gly Ile Arg Ala Met Thr Ser Leu Tyr Gly Glr.

(А явликаў — Ууч — Стана

	•				Ê					1(25	
5	Leu	Alā	Gly	Leu 20	Lys	Glu	Leu	Glγ	Leu 25	Leu	Asp	Суѕ	Хāā	Ser 30	Tyr	Il€
S	Thr	Gly	Ala 35	Ser	Gly	Ser	Thr	Trp 4(Ala	L€u	Ala	Asn	Leu 45	Тут	Lys	Asp
10	Fro	Glu 50	Trp	Ser	Gln	Lys	Asp 55	Leu	Ala	Gly	Pro	Thr 60	Glu	Leu	Leu	Lys
	Thr 65	Glr	Val	Thr	Lys	Asn 70	Lys	Leu	Gly	Val	Leu 75	Ala	Prc	Ser	Glr.	Leu 8(
15	Gln	Arg	Tyr	Arg	Gln 85	Glu	Leu	Ala	Glu	Arg 9(Ala	Arg	Leu	Gly	Tyr 95	Pro
20	Ser	Сув	Ph€	Thr 100	Asn	Leu	T'rp	Ala	Leu 105	Ile	Asn	Glu	Ala	Leu 11(Leu	Н15
20	Asp	Glu	Fro 115	His	Asp	His	Lys	1eu 12(Ser	Asp	Gln	Arg	Glu 125	Ala	Leu	Sei
25	His	Gly 130	Gln	Asn	Pro	L€u	Pro 135	Il∈	Tyr	Суѕ	Ala	Leu 140	Asn	Thr	Lys	Gly
	Gln 14!	Ser	Leu	Thr	Thr	Phe 150	Glu	Ph∈	Gly	Glu	Trp 155	Cys	Gîu	Ph€	Ser	Prc 160
30	Tyr	Glu	Val	Gly	Phe 165	Prc	Lys	Tyr	Gly	Ala 170	Phe	Ile	Pro	Ser	Glu 175	Leu
35	Ph∈	Gly	Ser	Glu 180	Ph∈	Ph∈	Met	Gly	Gln 185	Leu	Met	Lys	Arg	Leu 190	Pro	Glu
52	Ser	Arg	Il∈ 195	Сув	Phe	Leu	Glu	Gly 200	lle	Trp	Ser	Asn	Leu 205	Tyr	Ala	Ala
40	Asn	Leu 21(Gln	Asp	Ser	Leu	Tyr 215	Trp	Ala	Ser	Glu	Pro 220	Ser	Gln	Phe	Trp
	Asp 225	Arg	Trp	Val	Arg	Asn 230	Gln	Ala	Asn	Leu	Asp 235	Lys	Glu	Gln	Val	Pro 240
45	Leu	Leu	Lys	Ile	Glu 245	Glu	Prc	Pro	Ser	Thr 250	Ala	Gly	Arg	Ile	Ala 25t	Glu
50	Pł.€:	Phe	Thr	Asp 260	Leu	Leu	Thr	Trp	Arg 261	Pro	l.eu	Ala	Gln	Ala 270	Thr	His
30	Asn	Phe	Leu 275	Arg	Gly	Leu	His	Ph∈ 280	His	Lys	Asp	Tyr	Phe 285	Gln	His	Pro
55	His	Phe 29(Ser	Thr	Trp	Lys	Ala 295	Thr	Thr	Leu	Asp	Gly 300	Leu	Pro	Asn	Gln
	Leu 305	Thr	Pro	Ser	Glu	Pro 310	His	Leu	Cys	Leu	Leu 315	Asp	Val	Gly	Tyr	Leu 320
60	Ile	Asn	Thr	Ser	Cys	Leu	Pro	Leu	Leu	Gln	Pro	Thr	Arg	Asp	Val	Asp

PCT/US98/04482

W Exhaust A.

					325					330					335	
5	Leu	Il ϵ	Leu	Ser 340	Leu	Asp	Тут	Asn	Leu 345	His	Gly	Ala	Phe	Gln 350	Glr.	Lei
J	Gln	Leu	Leu 355	Gly	Arg	Phe	Cys	Gln 360	Glu	Gln	Gly	lle	Pro 365	Phe	Pro	Pro
10	Il€	Ser 370	Pro	Ser	Pro	Glu	Glu 375	Gln	Leu	Gln	Pro	Arg 380	Glu	Cys	His	Th
	Phe 385	Ser	Asp	Pro	Thr	Cys 390	Pro	Gly	Ala	Pro	Ala 395	Val	Leu	His	Phe	Pro 400
15	Leu	Val	Ser	Asp	Ser 405	Phe	Arg	Glu	Тут	Ser 410	Ala	Pro	Gly	Val	Arg 415	Arg
20	Thr	Pro	Glu	Glu 420	Ala	Ala	Ala	Gly	Glu 425	Val	Asn	Leu	Ser	Ser 430	Ser	Asp
20	Ser	Pro	Tyr 435	His	Tyr	Thr	Lys	Val 44(Thr	Tyr	Ser	Gln	Glu 445	Asp	Val	Asp
25	Lys	Leu 450	Leu	His	Leu	Thr	His 4 55	Tyr	Asn	Val	Cys	Asn 460	Asn	Gln	Glu	Glr
	Leu 465	Leu	Glu	Ala	Leu	Arg 47 0	Gln	Ala	Vāl	Gln	Arg 475	Arg	Arg	Gln	Arg	Arg 480
30	Pro	His	Xaa													
35	(0)	** \$ ***	OF.W.	77.ON	F05	CEO.			1.41							
55	(2)	11/17	ORMA'	SEQU.	ENCE	CHA	RACT.	ERIS'	TICS							
40			(xi)	(B) T D) T	YPE: OPOL	ami OGY:	no a lin	cid ear	aci EQ II		: 14	5:			
15	Met 1	Glu	Gly	Ala	Pro 5	Pro	Gly	Ser	L€u	Ala 10	Leu	Arg	Leu	Leu	Leu 15	Ph€
45	Val	Ala	Leu	Pro 20	Ala	Ser	Gly	Trp	Leu 25	Thr	Thr	Gly	Ala	Pro 30	Glu	Pro
50	Pro	Pro	Leu 35	Ser	Gly	Ala	Pro	Gln 40	Asp	Gly	Ile	Arg	Ile 45	Asn	Val	The
	Thr	Leu 50	Lys	Asp	Asp	Gly	Asp 55	Ile	Ser	Lys	Gln	Gln 60	Val	Val	Leu	Asr
55	Ile 65	Thr	Tyr	Glu	Ser	Gly 70	Gln	Val	Тут	Val	Asn 75	Asp	Leu	Pro	Val	Asr 80
60	Ser	Gly	Val	Thr	Arg 85	Ile	Ser	Cys	Gln	Thr 90	Leu	Ile	Val	Lys	Asn 9 <u>t</u>	Glu

	Asn	Leu	Glu	Asr. 100	Leu	Glu	Glu	Lys	Glu 105	Tyr	Phe	Gly	Ile	Val 11(Ser	Val
5	ЫĞ	Ile	Leu 115	Val	His	Glu	Trp	Prc 120	Met	Thr	Ser	Gly	Ser 125	Ser	Leu	Gln
	Leu	11e 130	Val	Ile	Gln	Glu	3lu 135	Val	Val	Glu	Il∈	Asp 14(Gly	Lys	Glr.	Val
10	Gln 145	Gln	Lys	Asp	Val	Thr 150	Blu	Ile	Asp	lle	Leu 155	Vāl	Lys	Asn	Arg	Gly 160
15	Val	Leu	Arg	His	Ser 165	Asr.	Tyr	Thr	Leu	Pro 170	Leu	Glu	Glu	Ser	Met 175	Leu
	Tyr	Ser	Ile	Ser 180	Arg	Asp	Ser	Asp	Il∈ 185	Leu	Phe	Thr	Leu	Prc 190	Asn	Leu
20	Ser	Lys	Lys 195	Glu	Ser	Val	Ser	Ser 200	Leu	Gln	Thr	Thr	Ser 20!	Gln	:Ax	Leu
	Il€	Arg 210	Asn	Val	Glu	Thr	Thr 215	Val	Asp	Glu	Asp	Val 220	Leu	Pro	Gly	Glr.
25	Val 225	Thr														
30	(2)	INF	ORMA!	TION	FOR	SEÇ	IDI	NO: 1	146:							
35				(A) L B) T D) T	CHA ENGT YPE: YOPOL E DE	H: 4 ami OGY:	5 am nc a lin	ino cid ear	acid		: 14	€:			
40	Met 1	Gly	Met	Gly	Ala 5	Phe	Gln	Ala	Phe	Phe 10	Trp	Val	Ile	Leu	Thr	Val
40	Ser	Asn	Val	Cys 20	Val	Leu	Phe	Lys	Met 25	Ser	Leu	Phe	Phe	Leu 3(Leu	Thr
45	Leu	Ile	Ser 35	Lys	Leu	His	Gly	Asp 40	Ala	Glu	Val	Cys	Xaa 45			
50	(2)	INF		SEÇU (ENCE (A) I (B) T	SEQ CHA ENGT TYPE:	FACT H: 1 ami	EFIS 32 a	TICS minc		ās					
55	Not	Cer				E DE Met								ጥኮኦ	G V	h -
	1	. 01	~+y	y						10		10	3		15	
60	Leu	Glv	Leu	Ala	Leu	Leu	Leu	Leu	Leu	Glv	Leu	Glv	Leu	Gly	Leu	Glu

				20					25					30		
•	Ala	Prc	Arg 35	Ala	Arg	Ph€	Pro	Pro 40	Arg	Pro	Leu	Pro	Arg 45	Pre	His	Pro
5	Ser	Ser 50	Gly	Ser	Cys	Pro	Pro 55	Thr	Lys	Phe	Gln	6(Cys	Arg	Thr	Ser	Gly
10	Leu 65	Суғ	Val	Pro	Leu	Thr 70	Trp	Arg	Суз	Asp	Arg 75	Thr	Trp	Thr	Ala	Ala 80
	Met	Ala	Ala	Met	Arg 85	Arg	Ser	Ala	Gly	Leu 90	Ser	His	Val	Pro	Arg 95	Lys
15	Gly	Asn	Ala	H1s 100	Arg	Pro	Leu	Ala	Ser 105	Pro	Ala	Pro	Ala	Pro 110	Ala	Ser
20	Val	Thr	Ala 115	Leu	Gly	Glu	Leu	Thr 120	Arg	Asn	Cys	Ala	Thr 125	Ala	Ala	Ala
20	Trp	Pro 130	Ala	Xaa												
25	(2)	INF	ORMA'	TION	FOR	SEQ	ID 1	NO: (148:							
30			(i) (xi)	(ENGT YPE : OPOL	H: 9 ami OGY:	2 am no a lin	ino cid ear	acid		: 14	٤:			
35	Met 1	Glu	Ala	Thr	Leu 5	Glu	Gln	His	Leu	Glu 10	Asp	Thr	Met	Lys	Asn 15	Pro
	Ser	Ile	Val	Gly 20	Vāl	Leu	Cys	Thr	Asp 25	Ser	Gln	Gly	Leu	Asn 30	Leu	Gly
40	Cys	Arg	Gly 35	Thr	Leu	Ser	Asp	Glu 40	His	Ala	Gly	Val	Ile 45	Ser	Val	Leu
45	Ala	Gln 50		Ala	Ala	Lys	Leu 55	Thr	Ser	Asp	Pro	Thr 60	Asp	Ile	Pro	Val
,,	Val 65	Cys	Leu	Glu	Ser	Asp 70	Asn	Gly	Asn	Ile	Met 75	Ile	Gln	Lys	His	Asp 80
50	Gly	Ile	Thr	Val	Ala 85	Val	His	Lys	Met	Ala 90	Ser	Xaa				
55	(2)	INF		SEQU (FOR ENCE A) L	CHA ENGT	RACT H: 1	ERIS 65 a	TICS mino		å r					
60			(xi)	(B) T D) T UENC	CPOL	OGY:	lin	ear	EQ I	D NO	: 14	9 :			

	Met 1	Glu	Prc	Leu	Arg £	Leu	Leu	Il€	Leu	Leu 10	Phe	Val	Thr	Glu	Leu 15	Ser
5	Gly	Ala	His	Asn 20	Thr	Thr	Val	Phe	Gln 25	Gly	Val	Ala	Gly	Gln 30	Ser	Leu
10	Gln	Vāl	Ser 35	Cys	Prc	Tyr	Asp	Ser 40	Met	Lys	His	Trp	Gly 4t	Arg	Arg	Lys
10	Ala	Trp 50	Cys	Arg	Gln	Leu	Gly 55	Glu	Lys	Gly	Pro	Cys 60	Gln	Arg	Val	Val
15	Ser 65	Thr	His	Asn	Leu	Trp 70	Leu	Leu	Ser	Phe	Leu 75	Arg	Arg	Trp	Asn	Gly 80
	Ser	Thr	Alā	Il∈	Thr 85	Asp	Asp	Thr	Leu	Gly 90	Gly	Thr	Leu	Thr	Ile 95	Thr
20	Leu	Arg	Asn	Leu 100	Gln	Pro	His	Asp	Ala 105	Gly	Leu	Tyr	Gln	Cys 110	Gln	Ser
25	Leu	His	Gly 115	Ser	Glu	Ala	Asp	Thr 120	Leu	Arg	Lys	Val	Leu 125	Val	Glu	Val
23	Leu	Ala 130	Asp	Pro	Leu	Asp	His 135	Arg	Asp	Ala	Gly	Asp 140	Leu	Trp	Phe	Pro
30	Gly 145		Ser	Glu	Ser	Phe 150	Glu	Asp	Ala	His	Vāl 155	Glu	His	Ser	Ile	Se: 160
	Arg	Ser	Ser	Ser	Xaa 16:											
35																
	(2)	INF	ORMA'													
40			(1)		A) I (B) I	ENGI	H: 1	39 a	mino		ā:					
			(xi)	SŁÇ	(D) I					EQ I	D NO	: 15	C :			
45	Met 1		Ser	Leu	Thr 5	Asp	Thr	Gln	Lys	Ile 10	Gly	Met	Gly	Leu	Thr 15	Gly
50	Ph∈	Gly	Val	Fh∈ 20	Phe	Leu	Ph∈	Phe	Gly 25	Met	lle	Leu	Phe	Phe 30	Asp	Lys
30	Ala	. Leu	Leu 35		lle	Gly	Asn	Val 40		Phe	Val	Ala	Gly 4º	Leu	Ala	Ph∈
55	Val	. Ile 50		' Leu	Glu	Arg	Thr 55		: Arg	: Phe	Phe	Phe 6(Glm	Lys	His	Lys
	Met 65	_	: Ala	Thr	Gly	Phe 7(Phe	Leu	Gly	· Gly	Val		Val	Val	Leu	Ile 80
60	Gly	/ Trp	Pro	Leu	· Ile	Gly	Met	Ile	Phe	e Glu	llì∈	Tyr	Gly	Ph€	Phe	Leu

					6 <i>r</i>					90					95	
	Leu	Phe	Arg	Glу 100	Phe	Ph€	Pro	Val	Val 105	Val	Gly	Phe	Ile	Arg 110	Arg	Val
5	Pro	Val	Leu 115	Gly	Ser	Leu	Leu	Asn 120	Leu	Pro	Gly	Ile	Arg 125	Ser	Phe	Va.
10	Asp	Lys 130	Val	Gly	Glu	Ser	Asn 135	Asn	Met	Val	Xaā					
15	(2)			(CHAI ENGT YPE:	RACT H: 5 ami	ERIS 8 am no a	TICS ino cid		l S					
20			(xi)	SEQ						EQ I	D NO	: 15	1:			
	Met 1	Ser	Ala	Pro	Gln £	Thr	Arg	Ile	Ser	Arg 10	Ala	Leu	Val	Leu	Leu 15	Ph€
25	Leu	Ala	Pro	Thr 20	Leu	Leu	Ser	Leu	Gly 25	His	Gly	Ile	His	Pro 30	Ile	Asn
20	Thr	Ala	Thr 35	Pro	Tyr	Xaa	Thr	Asp 40	Gln	Ala	Lys	Leu	Ala 45	Pro	Gly	Thr
30	Lys	Glu 50		Asn	His	Asp	Gln 55		Val	Thi						
35	(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:	152:							
40					(A) I (E) 7 (D) 7	LENGT TYPE : TOPOI	TH: 4 : am: LOGY	48 ar ino a : lir	mino acid near	acio): 15	52:			
45	Met 1		e Arg	lys	: Leu £	His	Lys	: Ile	e Ile	Val		Ser	Pro	Arg	Val 15	Il€
	Val	Leu	ı Lev	ı Asr 20		Ph∈	Phe	Phe	25 25		s Ala	Lys	: Phe	Val		Tyr
50	Ile	Phe	e Val		e His	; Val	. Lei	ı Asr 4(7 Sei	c Ile	e Ser	Ty:	Pro	Val	Xaa
55																
	(2)	INI	FORM	OITA	1 FO	R SE() ID	NO:	153	:						
60			(i)	SEQ'	UENC:	E CH	ARAC	TERI	STIC	S:						

 $\{(t_1, t_2, \dots, t_{n-1})_{n \in \mathbb{N}}\} = \{(\omega_1, \omega_2, \omega_3)_{n \in \mathbb{N}}\}$

```
(A) LENGTH: 42 amino acids
                  (E) TYPE: amino acid
                  (D) TOPCLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 153:
5
     Met Leu Leu Ash Gln His Phe Lys Ile Phe Gly Ser Leu Ile His Met
     Asr. Leu Leu Phe Ala Leu Ile Ser Leu Gly Ser Ser Asn Leu Ser Gly
10
        20 25
     Val Gln Phe Cys Cys Glu Thr Val Gln Xaa
            35
                              40
15
     (2) INFORMATION FOR SEQ ID NO: 154:
            (i) SEQUENCE CHARACTERISTICS:
20
                 (A) LENGTH: 72 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 154:
25
     Met Leu Ser Leu Ser Phe Leu Leu Arg Arg Val Leu Phe Leu Gly Phe
      1 5
     Leu Glm Ala Ser Val Gly Glu Lys Lys Ser Leu Arg Xaa Leu Asm Ty:
30
     Ser Val Pro His Pro Met Leu Xaa His Pro Pro Pro Asp Thr Ala Gl:.
                    40
     Val Pro Pro Arg Leu Glu Arg Ser Leu Leu Gln Gln Glu Leu Trp Thi
35
     Pro Gly Pro His His Ser Asn Ile
40
     (2) INFORMATION FOR SEQ ID NO: 155:
            (i) SEQUENCE CHARACTERISTICS:
45
                  (A) LENGTH: 106 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linea:
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 155:
      Met Gln Pro Leu Asn Phe Ser Ser Thr Glu Cys Ser Ser Phe Ser Pro
50
      Pro Thr Thr Val Ile Leu Leu Ile Leu Leu Cys Phe Glu Gly Leu Leu
55
      Phe Leu Ile Phe Thr Ser Val Met Phe Gly Thr Gln Val His Ser Ile
        35 40
      Cys Thr Asp Glu Thr Gly Ile Glu Gln Leu Lys Lys Glu Glu Arg Arg
       50
                55
                                      60
60
```

	Trp €5	Ala	Lys	Lys	Thr	Lys 70	Trp	Met	Asn	Met	Lys 75	Ala	Val	Phe	Gly	His 80
5	Pro	Fhe	Ser	Leu	Gly 85	Trp	Ala	Ser	Prc	Phe 90	Ala	Thr	Pro	qaA	Gîn 95	Gly
10	Lys	Ala	Asp	Pro 100	Tyr	Gln	Tyr	Val	Val 105	Xaa						
15	(2)	INF	(i)	(ENCE A) L B) T D) T	CHA ENGT YPE : OPOL	RACT H: 2 ami OGY:	ERIS 9 am no a	TICS nino ncid near	acić		·: 15	6 :			
20	Met 1	Syr	Thr	Asn	His 5	Phe	Asn	Leu	īyr	Leu 10	Lys	Tyr	Ile	Leu	Leu 15	${\tt Il}\epsilon$
25	Ile	Leu	Ile	Leu 20	Asn	Met	Thr	Asn	Ser 25	Ser	Ser	Arg	Тут			
30	(2)	INF		(ENCE (A) I (B) T	CHA LENGT	RACI H: !	ERIS	TICS nino acid		īs					
35	Met	Ler		SEÇ Leu	UENC	E DE	SCRI	PTIC	N: S					Phe	Thr	Ph∈
40	5				5 Asn					10					15 Phe	Leu
45			35 Trp	Tyr			Met	1 Leu 4(Pro	Pro	Val	Asn 45	Pro	Pro	Va.
50	(2)	INI	FORMA	MOITA	FOF	SE() ID	NO:	158:							
55					(A) 1 (B) ' (D) '	LENG IYPE ICPO	TH: : am LOGY	75 ai ino a : li:	mino acid near	aci		D: 15	58 :			
60	Met	Ty:	r Ala	a Val	. Тут 5		. Glr	ı Lei	a Ala	Glr 10		ı Thi	. Lei	ı Met	Val	Thr

need to yet designed

```
Leu Leu Ala Pro Ile Leu Pro Asp Glu Gln Ser Glu Val Phe Glu Ala
     Leu Ser Ash Leu Fro Lys Val Thr Trp Leu Gly Ser Ash Ser Fro Ser
         35 40
     Ser Glu Met Pro Glu Pro Gly Arg Phe Val Ile Val His His Gln Leu
10
     Ser Ala Ala Ser His Ser Ser Ser Gln Leu Ala
      €5 70
15
     (2) INFORMATION FOR SEC ID NO: 159:
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 81 amino acids
20
                  (B) TYPE: aminc acid
                 (D) TOPOLOGY: linear
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 159:
     Met Trp Pro Pro Leu Leu Leu Leu Leu Leu Leu Pro Ala Ala Pro
25
     Val Pro Thr Ala Lys Ala Ala Pro His Pro Asp Ala Asn Thr Gln Glu
30
     Gly Leu Gln Asn Leu Leu Gln Gly Val Gly Ala Gly Gly Asp Gly Glu
           35 4( 45
     Leu Arg Ala Asp Ser His Leu Ala Pro Gly Ser Gly Cys Ile Asp Gly
                 5 5
35
     Ala Val Val Ala Thr Arg Pro Glu Ser Arg Gly Gly Arg Pro Ala Val
      65 76 75
     Pro
40
      (2) INFORMATION FOR SEQ ID NO: 160:
45
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 139 amino acids
                  (B) TYPE: amino acid
                  (D) TOPCLOGY: linear
50
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16(:
     Met Lys Phe Thr Thr Leu Leu Phe Leu Ala Ala Val Ala Gly Ala Leu
                           10 15
      Val Tyr Ala Glu Asp Ala Ser Ser Asp Ser Thr Gly Ala Asp Pro Ala
55
      Glm Glu Ala Gly Thr Ser Lys Fro Asm Glu Glu Ile Ser Gly Pro Ala
                   4(
60
```

PCT/US98/04482

The Control of the Control

	Glu	Pro 50	Ala	Ser	Pro	Pro	Glu 55	Thr	Thr	Thr	Thr	Ala 60	Gln	Glu	Thr	Ser
5	Ala 65	Ala	Ala	Val	Gln	Gly 7(Thr	Ala	Lys	Val	Thr 75	Ser	Ser	Arg	Gln	Glน 80
	Leu	Asn	Prc	Leu	E E E	Ser	11ϵ	Val	Glu	Lys 90	Ser	Ile	Leu	Leu	Thr 95	Glu
10	Gln	Ala	Leu	Ala 100	Lys	Ala	Gly	Lys	Gly 105	Met	His	Gly	Gly	Val 110	Pro	Gly
15	Gly	Lys	Gln 115	Phe	Ile	Glu	Asn.	Gly 120	Ser	Glu	Phe	Ala	Gln 125	Lys	Leu	Leu
10	Lys	Lys 130	Phe	Ser	Leu	Leu	Lys 135	Pro	Trp	Ala	Xaa					
20	(2)	INFO	orma:	rion	FOR	SEQ	ID I	NO: I	161:							
25			, ,	(A) L B) T D) T	ENGT YPE: OPOL	H: 1 ami OGY:	ERIS' 78 a no a lin PTIO	mino cid ear	aci		: 16	1:			
30	Met 1	Leu	Gly	Cys	Gly 5	Il€	Pro	Ala	Leu	Gly 10	Leu	Leu	Leu	Leu	Leu 15	Gln
	Gly	Ser	Ala	Asp 20	Gly	Asn	Gly	Ile	Gln 25	Gly	Phe	Phe	Tyr	Pro 3(Trp	Ser
35	Cys	Glu	Gly 35	Asp	Ile	Trp	Asp	Arg 40	Glu	Ser	Cys	Gly	Gly 45	Gln	Ala	Ala
40	Ile	Asp 50		Pro	Asn	Leu	Cys 5:	Leu	Arg	Leu	Arg	Cys 6(Cys	Tyr	Arg	Asn
	Gly 65	Val	Cys	Tyr	His	Gln 70	Arg	Prc	Asp	Glu	Asn 75	Val	Arg	Arg	Lys	His 80
45					8£			Cys		90					95	
				100				Ala	105					110		
50			115					Суs 120					125			
55		130					135					140				
	145					150		: Ser			155					160
60	Gly	Glu	Gly	Thr	Glu 165		Gly	Glu,	Glu	Thr 170		Gly	Glu	Glu	. Glu 175	

Asp Xaa

5

(2) INFORMATION FOR SEQ ID NO: 162:

(i) SEQUENCE CHARACTERISTICS:

10 (A) LEWSTH: 72 amino acids

(E) TYPE: amino acid

(D) TOPULOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 162:

15 Met Glu Ala Val Fhe Thr Val Phe Phe Phe Val Val Val Leu Phe Leu $_1$

Lys Asn Thr Glu Gly Ala Lys Leu Phe Cys Thr Leu Tyr Pro Ala Ala

20

Ser Ser Gly Gln Ser Gln Gly Pro Gly Leu Glu Lys Pro Asp Ser Gln 35 40 45

Glu Cys Ile Ile Asp Pro Cys Ser Tyr Pro Ile Ala Leu Gly Ala Gly 25 50 55 60

Thr Glu Pro Gly Cys Lys Ile Xaa 65 7(

30

35

(2) INFORMATION FOR SEQ ID NO: 163:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 67 amino acids

(B) TYPE: amino acid

(D) TOFOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 163:

40 Met Trp Phe Tyr Phe Leu Ser Val Ser Phe Pro Leu Leu Pro Val Xaa 1 1 15

Ala Pro Xaa Pro Pro Pro Ala Pro Thr Thr Leu Cys Leu Leu Phe 20 25 30

45 Leu Gly Xaa Leu Tyr Asn Ser Thr Cys Ile His Cys Val His Thr Thr 35 40 45

Ser Xaa Thr Gln Asn Frc Thr Ala Asn Thr Leu Lys Lys Lys Lys 50 $_{\rm 50}$ 60 $_{\rm 50}$

Asn Trp Gly

55

(2) INFORMATION FOR SEQ ID NO: 164:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 155 amino acids

THE THE WALLEY

						ipe: CPOL										
			(xi)	SEÇ	JENCI	E DES	SCRII	PTIO	N: S1	EQ I	D NO	: 164	:			
5	Met 1	Gly	Phe	Gly	Ala £	Thr	Leu	Ala	Val	Gly 10	Leu	Thr	Ile	Phe	Val 15	Leu
10	Ser	Val	Val	Thr 20	Il€	Ile	Ile	Cys	Phe 25	Thr	CÀE	Ser	Cys	Cys 30	Cys	Leu
10	Tyr	Lys	Thr 35	Cys	Arg	Arg	Pro	Arg 40	Pro	Val	Val	Thr	Thr 45	Thr	Thr	Ser
15	Thr	Thr 50	Val	Val	His	Ala	Pro 55	Tyr	Pro	Gln	Pro	Pro 60	Ser	Val	Pro	Prc
	Ser 65	Tyr	Fro	Gly	Pro	Ser 70	Tyr	Gln	Gly	Tyr	His 75	Thr	Met	Pro	Pro	Gln 80
20	Pro	Gly	Met	Pro	Ala Et	Ala	Pro	Tyr	Pro	Met 90	Gln	Tyr	Pro	Pro	Pro 95	Tyr
25	Pro	Ala	Gln	Pro 100	Met	Gly	Pro	Pro	Ala 105	Tyr	His	Glu	Thr	Leu 110	Ala	Gly
	Glu	Gln	Pro 115	Arg	Pro	Thr	Pro	Pro 120	Ala	Ser	Leu	Leu	Thr 125	Thr	Arg	Pro
30	Thr	Trp 130	Met	Prc	Arg	Arg	Arg 135	Pro	Ser	Glu	His	Ser 140	Leu	Ala	Ser	Leu
	Ala 145	Ala	Thr	Trp	Leu	Cys 150	Cys	Val	Cys	Ala	Xaa 15!					
35	(2)	INF	orma'	NOIT	FOR	SEQ	ID 1	NO: í	165:							
40				SEÇU	ENCE	CHA ENGT	RACT	ERIS'	rics		ā:					
			(xi)	(B) T D) T	YPE: OPOL E DE	ami OGY:	no a lin	cid ear			: 16	5 :			
45	Met 1	Ile	Ile	Leu	Val	Phe	Ile	Ala	Phe	Phe	Ile	Pro	Leu	Gln	Lys 15	Thr
	Ile	Gly	Lys	Ile 20	Ala	Thr	Cys	Leu	Glu 25	Leu	Arg	Ser	Ala	Ala 30	Leu	Gln
50	Ser	Thr	Gln 35		Gln	Glu	Glu	Phe 40	Lys	Leu	Glu	Asp	Leu 45	Lys	Lys	Leu
55	Glu	Pro 50		Leu	Lys	Asn	Ile 55	Leu	Thr	Tyr	Asn	Lys 6(Glu	Phe	Pro	Phe
	Asp 65	Va]	Gln	Pro	Val	Pro 70	Leu	Arg	Arg	Ile	Leu 71	Ala	Pro	Gly	Glu	Glu 80
60	Glu	Asn	Leu	Glu	Phe	Glu	Glu	Asp	Glu	Glu	Glu	Gly	Gly	Ala	Gly	Ala

the supplied whateh

```
3.3
                                     9(
                                                        9:
     Gly Leu Leu Ile Leu Ser Cys Xaa
        10(
5
     (2) INFORMATION FOR SEQ ID NO: 166:
10
          (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 81 amanc acids
                  (E) TYPE: amino acid
                  (D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 166:
15
     Met Ala Gly Thr Met Val Ile Val Val Val Val Val Val Gly Glu Val
     Val Val Glu Ala Glu Val Val Val Gln Ala Arg Glu Glu Ala Gly Glu
20
     Glu Glu Gly Ala Arg Ile Ile Thr Lys Gly Val Asn Leu Asn Ser Ile
25
     Ser Ser Met Glu Val Ile Ser Ile Ile Ile Leu Asp Leu Asp Arg Glu
                           55
                                    60
     Asp Ile Thr Leu Val Glu Ala Thr Glu Pro Tyr Ile Leu Leu Glu Leu
                                           77.5
                  7(
30
     Lys
35
     (2) INFORMATION FOR SEC ID NO: 167:
            (i) SECUENCE CHARACTERISTICS:
                  (A) LENGTH: 93 amino acids
40
                   (E) TYPE: amino acid
                  (D) TOPOLOGY: linear
            (xi) SEÇUENCE DESCRIPTION: SEÇ ID NO: 167:
     Met Ser Phe Ser Phe Ile Ile Phe Leu Leu Leu Val Cys Gln Glu Ile
45
      1 5
      Thr Phe Cys Met Ser Tyr Gly Asp Ala Val Ash Cys Phe Ser Glu Cys
50
      Phe Ser Asn Leu Gln Thr Ile Tyr Ile Ser Cys Leu Gln His Ala Val
             35 40
      Cys Lys His Ser Val Ile Trp Ser Ile Gln Leu Phe Val Arg Ala Leu
                 55
55
      Pro Ile Ser Lys Cys Ala Glu Leu Ser Ile Asp Gly Ile Phe Arg Se:
                                           7:
      Phe His Glu Asn Trp Lys Cys Ser Trp Val Ala Pro Thr
60
                    2.8
```

5	(2)	INFO	RMAT	NOI	FOR	SEQ	ID N	10: 1	68:							
5			(i) S	(1	A) L1 B) T	ENGTI YPE :	H: 5 amin	ERIST 8 am: no ac line	ino a							
10			(xi)							EQ II	NO:	: 168	ā:			
	Met 1	Gly	Trp	Ser	Ala 5	Gly	Leu	Leu	P'n€	Leu 1(Leu	Ile	Leu	Tyr	Leu 15	Prc
15	Val	Pro	Gly	Trp 20	Met	Glu	Arg	Glu	Asp 25	Gly	Glu	Thr	Gly	His 30	Leu	Sei
20	Pro	Gln	Ala 35	Pro	Gly	Arg	Glu	Туг 40	Arg	Gly	Phe	Tyr	Ser 45	Val	Pro	Prc
20	Asp	Tyr 50	Val	Trp	Leu	Arg	Asp 55	Ser	Prc	Χāε						
25	(2)	INF	ORMAT	NOLT	FOR	SEQ	ID I	NO: 3	169:							
30			(i) s (xi)	(A) L B) T D) T	ENGT YPE : CPOL	H: 2 ami OGY:	ERIS 32 a no a lin PTIO	mino cid ea:	aci		: 16	9:			
35	Met 1	Ala	Thr	Leu	Trp 5	Gly	Gly	Leu	Leu	Arg 10	Leu	Gly	Ser	Leu	Leu 15	Sei
	Leu	Ser	Cys	Leu 20	Ala	Leu	Ser	Val	Leu 25	Leu	Leu	Ala	His	Cys 30	Gln	Thi
40	Pro	Pro	Arg 35	Ile	Ser	Arg	Met	Ser 40	Asp	Val	Asn	Val	Ser 45	Ala	Leu	Pro
45	Ile	Lys 50	Lys	Asn	Ser	Gly	His 55	Ile	Tyr	Asn	Lys	Asn 60	Ile	Ser	Gln	Lys
43	Asp 65	Cys	Asp	Cys	Leu	His 70	Val	Val	Glu	Pro	Met 75	Pro	Val	Arg	Gly	Pro 80
50	Asp	Val	Glu	Ala	Tyr 85	Cys	Leu	Arg	Сує	Glu 9(Cys	Lys	Tyr	Glu	Glu 95	Arç
	Ser	Ser	Val	Thr 10(Ile	Lys	Val	Thr	Ile 105	Ile	Ile	Tyr	Leu	Ser 110	Ile	Leu
55	Gly	Leu	Leu 115	Leu	Leu	Tyr	Met	Val 120	Tyr	Leu	Thr	Leu	Val 125		Pro	Il€
60	Leu	Lys 130	Arg	Arg	Leu	Phe	Gly 135		Ala	Gln	Leu	11e	Gln	Ser	Asp	Ası

	Asp 145	li€	G _* ; _'	ASP	His	150	F2 C	PL€	Aic	ASN	155	HIS	ASP	Vāl	Leu	16(
5	Arç	192	Arg	Sei	Arg 165	Ala	Asn	Val	Leu	Asn 17(Lys	Val	Glu	Tyr	Gly 17t	Tha
	Alā	Ala	Leu	Glu 18(Ala	Ser	Ser	Pro	Arg 181	Ala	Alā	Lys	Ser	Leu 190	Ser	Leu
10	Thr	Gly	Met 195	Leu	Ser	Ser	Ala	Asr. 20(Trp	Gly	Ile	Glu	Phe 205	Lys	Val	Thr
15	Arg	Lys 21(Lys	alD	Ala	Asp	Asn 215	Trp	Lys	Gly	Thr	Asp 220	Trp	Val	Leu	Leu
	Gly 225	Ph€	Il€	Leu	Ile	Pro 230	Cys	Xãé								
20	(2)	INF	orma'	IION	FOR	SEQ	ID I	10: :	176 :							
25			(i)	(A) L		H: 7	2 am	TICE inc		Æ					
			(xi)			OPOL E DE			eai N: S	EÇ I	D N O	: 17	0:			
30	7				5				Ser	10					1 5	
				20					Arg 2t					30		
35			35					4(Gly				45			
40		5(Arg Ile			5.5		Val	Ala	Val	Cys	Сув	Ile	Vā.	Ke¹
15	65					70										
45	(2)	INF		TION												
50					(A) I (B) T (D) T	.ENGT TYPE : TOPOI	H: 6 ami OGY:	5 am no a lir		acid): 1 7	1:			
55	Met 1	Gly	Thr	Phe	Ser 5		Ser	Leu	Phe	Gly 10		Met	Gly	Val	Ala 15	
	Gly	Met	Asn	Leu 20		Ser	Ser	Leu	Glu 25		. Asp	His	Arg	: Il∈ 3(Phe	Irr
60	Leu	Ιlε	Thr	Glv	· Ile	Met	Phe	Met	Gly	Ser	Gly	Leu	Ile	Trp	Arg	Arç

WO 98/39446

			35					4(45			
5	Leu Val	Leu 50		Phe	Leu	Gly	Arg 55	Gln	Leu	Glu	Ala	Pro 60	Leu	Pro	Pro	Met
10	(2)	INF	ORMA'	TION	FOR	SEÇ	ID	NO:	172 :							
15				((A) I (B) T (D) T	ENGT YPE : YPOL	H: 7 ami OGY:	75 am no a lin	uno cic mear	acid		: 17	2:			
20	Met 1	Tyr	Lys	Gly	Lys 5	Leu	Val	Ile	Val	Leu 10	Ile	Leu	Leu	Leu	Leu 15	Pro
	Ser	His	Ph€	Met 20	Phe	Leu	Thr	Gln	Cys 25	Lys	Glu	Ile	Lys	His 3(Asn	Leu
25	Lys	Lys	Asn 35	Met	Ser	Leu	Leu	Leu 4(Phe	Thr	Ile	Lys	Ser 45	Trp	Leu	Tyr
30	Ser	Ala 50	Ser	Leu	Gly	Il∈	Leu 55	∵yr	Asn	Trp	Gln	His 60	Leu	Thr	Ala	Glr.
50	Val 65	Asp	Gln	Cys	Thr	Ser 7(Leu	Ile	Leu	Ile	Ηiε 75					
35	(2)	INF	ORMA!	rion	FOR	SEQ	ID I	NO: :	173:							
40				(A) L B) T D) T	ENGT YPE: OPOL	H: 3 ami OGY:	34 a no a lin	mino cić ear	aci		: 17	3:			
45	Met 3	Val	Gly	His	Glu 5	Met	Ala	Ser	Xaa	Ser 10	Ser	Asn	Thr	Ser	Leu 15	Prc
	Phe	Ser	Asn	Met 20	Gly	Asn	Pro	Met	Asn 25	Thr	Thr	Gln	Leu	Gly 30	Lys	Ser
50	Leu	Phe	Gln 35	Trp	Gln	Val	Glu	Gln 40	Glu	Glu	Ser	Lys	Leu 45	Ala	Asn	Ile
55	Ser	Gln 50	Asp	Gln	Phe	Leu	Ser 55	Lys	Asp	Ala	Asp	Gly 60	Asp	Thr	Phe	Leu
	His 65	Ile	Ala	Val	Ala	Gln 7(Gly	Arg	Arg	Ala	Leu 75	Ser	Tyr	Val	Leu	Ala 80
60	Arg	Lys	Met	Asn	Ala 85	Leu	His	Met	Leu	Asp 90	Ile	Lys	Glu	His	Asn 95	GЈУ

ENGINE OF SURE

	Gln	Ser	Ala	Phe	Gln	Vāl	Ala	Val	Ala 105	Ala	Asn	Gln	His	Leu 11(Ile	Val
5	Gln	Asp	Leu 115	Val	Asn	Ile	Gly	Ala 120	Gln	Val	Asn	Thr	Thr 125	Asp	Cys	Trp
10	Gly	Arg 130	Thr	Frc	Leu	His	Val 135	€£	Ala	Glu	Lys	Gly 14(His	Ser	Gln	Val
10	Leu 145	Gln	Ala	Ile	Glr.	Lys 150	Gly	Ala	Val	Gly	Ser 155	Asn	Gln	Phe	Val	Asp 160
15	Leu	Glu	Ala	Thr	Asn 165	Тут	Asp	Gly	Leu	Thr 170	Pro	Leu	His	CÀE	Ala 175	Val
	Il€	Ala	His	Asn 180	Ala	Vāl	Val	His	Glu 185	Leu	Gln	Arg	Asn	Gln 190	Gln	Prc
20	His	Ser	Pro 195	Glu	Vāl	Gln	Glu	Leu 200	Leu	Leu	Lys	Asn	Lys 20t	Ser	Leu	Val
25	Asp	Thr 210		Lys	Cys	Leu	11e 215		Met	Gly	Ala	Ala 22(Val	Glu	Ala	Lys
23	Asp 225	Arg	Lys	Ser	Gly	Arg 230		Ala	Leu	His	Leu 23!	Ala	Ala	Glu	Glu	Ala 240
30	Asn	Leu	Glu	Leu	11e 241	Arg	Leu	ı Ph∈	Leu	G1u 250		Pro	Ser	Cys	Leu 255	Ser
	Phe	Val	Asr	Ala 260		Ala	туг	: Asn	Gly 265		Thr	Ala	Leu	His 270		Ala
35	Ala	Ser	Let 275		Tyr	Arg	, Lei	280		Lev	a Asp	Ala	Val 285	Arg	Leu	Leu
40	Met	Arg 290		Gly	/ Ala	Asp	295		Thr	: Arg	Asr	1 Leu 30(. Asn	Glu	Glr.
40	Pro 305		His	s Lev	ı Val	310		o Gly	Pro	Va.	1 Gly 315		: Glr	ıle	e Arg	320
45	Iì€	Lei	ı Ly:	s Gly	y Lys 325		r Il	e Glr	n Glr	330	g Ala	e Pro	Pro	Ty:		
50	(2)	INI						NO:								
			(i)	SEÇ	(A) (B)	LENG TYPE	TH: I: an	TERI. 196 hino 7: li	amin acid	o ac	ids					
55			(xi) SE				IPTI			ID N	0: 1	74:			
		t As; 1	p Al	a Ar	g Tr	p Tr	p Al	a Va	l Va	1 Va 1		u Al	a Al	a Ph	e Pr 1	o Ser
60	•	. 01	20	- C1	v Gl	v Gl	u Th	r Pr	o G1	u Al	a Pr	o Pr	o G1	u Se	r Tr	p Thi

				2(25					30		
5	Gln	Leu	Trp 35	Ph∈	Ph∈	Arg	Phe	Val 40	Vāl	Asn	Ala	Alā	Gly 45	Tyr	Ala	Ser
-	Phe	Met 50	Val	Pro	Gly	Туг	Leu 55	Leu	Vāl	Gln	Tyr	Phe 60	Arg	Arg	Lys	Asr
10	Tyr 65	Leu	Glu	Thr	Gly	Arg 70	Gly	Leu	ርን፡ε	Phe	Pro 71	Leu	Vāl	Lys	Ala	Cys 80
	Val	Phe	Gly	Asn	Glu 85	Pro	Lys	Ala	Ser	Asp 90	Glu	Val	Pro	Leu	Ala 95	Pro
15	Arg	Thr	Glu	Ala 100	Ala	Glu	Thr	Thr	Pro 105	Met	Trp	Gln	Ala	Leu 110	Lys	Leu
20	Leu	Phe	Cys 115	Ala	Thr	Gly	Leu	Gln 120	Val	Ser	Tyr	Leu	Thr 125	Trp	Gly	Va]
20	Leu	Gln 130	Glu	Arg	Val	Met	Thr 135	Arg	Ser	Tyr	Gly	Ala 140	Thr	Ala	Thr	Ser
25	Pro 1 4 5	Gly	Glu	Arg	Phe	Thr 150	Asp	Ser	Gln	Phe	Leu 155	Val	Leu	Met	Asn	Arg 160
	Val	Leu	Ala	Leu	Ile 16!	Val	Ala	Gly	Leu	Ser 17(Cys	Val	Leu	Cys	Lys 175	Glr
30	Pro	Arg	His	Gly 180	Ala	Pro	Met	Tyr	Arg 185	Tyr	Ser	Phe	Cys	Gln 190	Pro	Va.
35	Gln	Cys	Ala 195	Xaē												
40	(2)	INF	(i)	SEQU () (ENCE A) L B) T D) T	CHA ENGT YPE:	RACT H: 2 ami OGY:	NO: SERISSON AND A LIN	TICS mino cid ear	aci		17	r			
45	Met	Ser						PTIO Gly		Ile				Thr	Leu	Lei
50	1 Leu	Leu	Leu	Thr 20	Leu	Leu	Ala	Phe	Ala 25	10 Gly	Tyr	Ser	Gly	Leu 30	15 Leu	Ala
	Gly	Val	Glu 35	Val	Ser	Ala	Gly	Ser 40	Pro	Pro	Ile	Arg	Asn 45	Val	Thr	Va.
55	Ala	Tyr 50	Lys	Phe	His	Met	Gly 55	Leu	Tyr	Gly	Glu	Thr 60	Gly	Arg	Leu	Phe
60	Thr 65	Glu	Ser	CAs	Ser	lle 70	Ser	Pro	Lys	Leu	Arg 7!	Ser	Ile	Ala	Val	Ty:

	lyr	Asp	Asn	Prc	His Et	M∈t	Vāl	Prc	Prc	Asp 9(Lys	Cys	Arg	Cys	Ala 91	Val
5	Gly	Sex	Il∈	Leu 10(Ser	Glu	Gly	Glu	Glu 105	Ser	Frc	Ser	Prc	Glu 110	Leu	Il€
	Asp	Leu	Tyr 115	Gln	Lys	Phe	Gly	Phe 120	Lys	Val	Ph€	Ser	Phe 125	Prc	Glu	Pro
10	Ser	His 130	Vāl	Val	Thr	Alā	Thr 135	Phe	Frc	Leu	Thr	Prc 14(Pro	Phe	Суѕ	Pro
15	Ile 145	Trp	Leu	Gly	Tyr	Frc 150	Pro	Cys	Pro	Ser	Cys 155	Leu	Gly	His	Leu	His 160
10	Gln	Gly	Ala	Glu	Ala 165	Val	Cys	Leu	Ser	Ser 170	Ala	Gly	Asp	Leu	Pro 175	Gly
20	Arg	Pro	Glu	Ser 18(Il€	Ser	CAs	Ala	His 185	Trp	His	Gly	Gln	Gly 190	Asp	Ph€
	Tyr	Val	Frc 195	Glu	Met	Lyε	Glu	Thr 200	Glu	Trp	Lys	Trp	Arg 205	Gly	Leu	Va.
25	Glu	Ala 210	Ile	Asp	Thr	Gln	Val 215	Asp	Gly	Thr	Gly	Ala 220	Asp	Thr	Met	Sei
30	Asp 225		Ser	Ser	Val	Ser 230	Leu	Glu	Val	Ser	Pro 23:	Gly	Ser	Arg	Glu	Th: 24(
	Ser	Ala	Ala	Thr	Leu 245	Ser	Pro	Gly	Ala	Ser 25(Ser	Arg	Gly	Trp	Asp 255	Asr
35	Gly	Asp	Thr	Arg 260		Glu	His	Ser	Xa a 265							
40	(2)	INF	ORMA	SEQU	JENCE (A) 1 (B) !	CHA LENGT	RACI TH: 1	ERIS 138 a ino a	TICS mind		Ιđε					
45			(xi)	SEÇ				: lir [PTIC		SEÇ I	D N	0: 17	∂€:			
	Met 1		Glr	ı Lev	Ph∈		Prc	Leu	Leu	Ala 10		Leu	ı Val	. Leu	Ala 15	Glr.
50	Ala	Pro	Ala	a Ala 20		ı Alā	Asp	Val	Leu 25		: Gly	/ Asp	Ser	Ser 30		ı Asp
55	Arg	g Ala	a Phe		g Val	. Arg	j Il∈	Ala 40		/ Asp	Ala	a Pro	Lev 45		Gly	/ Vāl
55	Lev	ı Gly 50		/ Ala	a Let	ı Thi	: Ile		Cys	s His	s Val	His 60		. Let	ı Arç	g Pro
60	Pro 65		o Sei	r Arg	g Arq	Ala و 70		. Let	ı Gly	/ Se:	7!		g Va.	l Lys	s Trp	1.4T c 38

	Ph€	L∈u	Ser	Arg	Gly 85	Arg	Glu	Ala	Glu	Val 90	Leu	Val	Ala	Arg	Gly 95	Val
5	Arg	Val	Lys	Val 100	Asn	Glu	Ala	Tyr	Arg 105	Phe	Arg	Val	Ala	Leu 110	Pro	Ala
10	Тут	Pro	Ala 115	Ser	Leu	Thr	Asp	Val 120	Ser	Pro	Gly	Ala	Glu 125	Arg	Ala	Ala
•	Pro	Gln 130	Arg	Leu	Arg	Tyr	Leu 135	Ser	Leu	Хағ						
15	(2)	INFO	ORMA'	noic	FOR	SEQ	IDI	NO: 3	177 :							
20				(ENCE A) L B) T D) T UENC	ENGT YPE : OPOL	H: 1 ami OGY:	79 a no a lin	mino cid ear	aci		: 17	7 :			
25	Met 1	Pro	Ala	Leu	Arg E	Pro	Ala	Leu	Leu	Trp	Alā	Leu	Leu	Ala	Leu 15	Trp
	Leu	Cys	Сув	Ala 20	Thr	Pro	Ala	His	Ala 25	Leu	Gln	Cys	Arg	Asp 30	Gly	Tyi
30	Glu	Pro	Cys 35	Vāl	Asn	Glu	Gly	Met 40	Cys	Val	Thr	Tyr	His 45	Asn	Gly	Thi
35	Gly	Tyr 50	Cys	Lys	Gly	Pro	Glu 55	Gly	Phe	Leu	Gly	Glu 60	Tyr	Cys	Gln	His
	Arg 65	Asp	Pro	C?.e	Glu	Lys 70	Asn	Arg	Cys	Gln	Asn 7t	Gly	Gly	Thr	Cys	Val 80
40	Ala	Gln	Ala	Me∙t	Leu 85	Gly	Lys	Ala	Thr	Cys 9(Arg	Cys	Ala	Ser	Gly 95	Ph∈
	Thr	Gly	Glu	Asp 10(Cys	Gln	Tyr	Ser	Thr 105	Ser	His	Pro	Cys	Phe 110	Val	Sez
45	Arg	Pro	Cys 115	Leu	Asn	Gly	Gly	Thr 120	Cys	His	Met	Leu	Ser 125	Arg	Asp	The
50	Tyr	Glu 130	Cys	Thr	Cys	Gln	Val 135		Phe	Thr	Gly	Lys 14(Glu	Cys	Gln	Trp
	Thr 145		Ala	Cys	Leu	Ser 150	His	Pro	Cys	Ala	Asn 15t	Gly	Ser	Thr	Cys	Thr 160
55	Thr	Val	Ala	Asn	His 165	Phe	Leu	Gln	Met	Pro 17(His	Arg	Leu	His	Arg 175	Ala
	Glu	Val	Xaε.													

	(2)	1NF(RMAI	.I C#1	FOR	SEÇ	ID N	:0: 1	.78 :							
5			(i) S (xi)	() () ()	A) LI E) T D) T	ENGT: YPE : OPOL(H: 1: amii DGY:	55 a nc a lin	minc cić ear	aci		: 178	::			
10	Met :	Thr	Arg	Gly	Gly 5	Prc	Glγ	Gly	Μģ	Frc 10	Gly	Leu	Pro	Gln	Pro 15	Fre
1.5	Prc	Leu	Leu	Leu 20	Leu	Leu	Leu	Leu	Prc 25	Leu	Leu	Leu	Val	Thr 30	Ala	Glu
15	Frc	Frc	35 778	Pro	Ala	Gly	Val	Tyr 4(፤ን።	Alā	Thr	Ala	Tyr 45	Trp	Met	Fro
20	Ala	Glu 5(Lys	Thr	Val	Gln	Val 55	Lyf	Asn	Vāl	Met	Asp 60	Lys	Asn	Gly	Ası
	Ala 65	Tyr	Gly	Ph∈	Tyr	Asn 70	Asn	Ser	Vāl	Lys	Thr 75	Thr	Gly	Trp	Gly	11¢ 80
25	Leu	Glu	lle	Arg	Ala 85	Gly	Tyr	Gly	Ser	Gln 90	Thr	Leu	Ser	Asn	Glu 9E	ĭl€
30	Ile	Met	Phe	Val 100	Ala	Gly	Ph€	Leu	Glu 10:	Gly	Tyr	Leu	Ile	Ala 11(Prc	Hii
50	Met	Asrı	Asp 115	His	Тут	Thr	Asn	L∈u 12(Tyr	Prc	Gln	Leu	11e 125	Thr	Lys	Pro
35	Ser	Il∈ 13(Met	Asp	Lys	Val	Gln 135	Asp	Phe	Met	Glu	Lys 140	Gln	Asp	Lys	Val
	Asp 145	Рτс	Glu	Lys	Tyr	Gln 150	Arg	Il€	Gln	qsA	Xãã 155					
40						950		•••	150							
45	(2)	INF		SEQU	ENCE (A) I (B) I (D) I	CHA LENGI TYPE:	RACT TH: 2 : ami	ERIS 295 a ino a : lir	TICS amino acid nea:	o aci): 17	9:			
50	Met 1	Leu	: Gln	. Gly	Pro		· Ser	Leu	ı Leu	Leu 1(. Ph ϵ	Leu	Ala	Ser	His
	Cys	: Cys	: Leu	Gly 20		Ala	. Arg	: Gly	r Leu 2'		e Leu	ı Phe	Gly	Glr 30		ASI
55	Phe	- Sez	: Tyr 35		: Arg	r Xaa	Asn	ı Cys		Pro	Ξle	e Pro	Val		. Leu	Glr
60	Leu	ı Cys	His	Gly	⁄ Il∈	Glu	n Tyr 55		. Asr	n Met	. Arg	Leu 60) Asr	ı Leu	ı Leu

	Gly 65	His	Glu	Thr	Met	Lys 7(Glu	Val	Leu	Glu	Gln 75	Ala	Gly	Ala	Trp	11€ 80
5	Frc	Leu	Val	Met	Lys 85	Gln	CAE	His	Pro	Asp 90	Thr	Lys	Lys	Phe	Leu ș <u>i</u>	СÀŧ
10	Ser	Leu	Ph∈	Ala 100	Pro	Vāl	Сля	Leu	Asp 105	Asp	Leu	Asp	Glu	Thr 11(Ile	Glr.
10	Pro	Cys	His 115	Ser	Leu	Сує	Val	Gln 120	Val	Lys	Asp	Arg	Сув 125	Ala	Pro	Val
15	Met	Ser 130	Ala	Phe	Gly	Ph∈	Pro 135	Trp	Pro	Asp	Met	Leu 140	Glu	Cys	Asp	Arg
	Phe 145	Pro	Gln	Asp	Asn	Asp 150	Leu	Cλε	Ile	Pro	Leu 155	Ala	Ser	Ser	Asp	H18 160
20	Leu	Leu	Pro	Ala	Thr 165	Glu	Glu	Ala	Pro	Lys 170	Val	Cys	Glu	Ala	Cys 175	Lys
25	Asn	Lys	Asn	Asp 180		Asp	Asn	. Asp	Ile 185	Met	Glu	Thr	Leu	Cys 19(Lys	Asn
23	Asp	Phe	Ala 195	Leu	Lys	Ile	Lys	Val 200	Lys	Glu	Ile	Thr	Tyr 205	Ile	Asn	Arg
30	Asp	Thr 210		Ile	Ile	Leu	Glu 215		Lys	Ser	Lys	Thr 220	Il€	Tyr	Lys	Leu
	Asn 22!	Gly	Val	Ser	Glu	Arg 230		Leu	Lys	Lys	Ser 235	· Val	Leu	Trp	Leu	Lys 240
35	Asp	Ser	Leu	ı Glm	Cys 245		- Cys	Glu	Glu	Met 250		Asp	lle	Asn	Ala 255	Pro
40	Tyr	Leu	ı Val	Met 260		Glr	. Lys	Glr.	Gly 265		/ Glu	Leu	ı Val	11∈ 27(Thr	Ser
40	Val	Lys	275		Glr.	Lys	Gl <u>y</u>	/ Glr 280		g Gli	ı Phe	E Lys	285	ı Ile	e Ser	Arç
45	Ser	7 Ile 290		j Lys	Leu	ı Glr	n Cys 295	-								
50	(2)) INI			UENCI (A) (B)	E CH LENG TYPE	ARAC TH:	NO: TERI 256 lino	STIC amin acid	S: o ac	riās					
55			(xi) SE				: li IPTI			ID N	0: 1	8C:			
		t Ar î	g Pr	c Al		a Le	u Ar	g Gl	y Al	a Le 1		u Gl	у Су:	s Lei	и Су 1	s Leu L
60	Al	a Le	u Le	u Cy	s Le	u Gl	y Gl	y Al	a As	p Ly	s Ar	g Le	u Ar	g As	p As	n His

Professional War Beautiful A

				2(2 :					3(
5	Glu	Trp	Lys 3t	Lys	beu	$Il\epsilon$	Met	Val 40	Glr.	His	Trp	Pro	Glu 4!	Thr	Val	Сує
3	Glu	Lys 50	Ile	Glr.	Asr.	Asp	Cys 5t	Αrg	Asp	Prc	Pro	Asp E(Tyr	Trp	Thr	Πe
10	His 65	Gly	Leu	Trp	Frc	Asp 70	Lys	Ser	Glu	Gly	Cys 75	Asn.	Ara	Ser	Trp	Prc 80
	Fh∈	Asn	Leu	Glu	Glu E£	Il€	Lys	Asp	Leu	Leu 90	Prc	Glu	Met	Arg	Ala 95	Туг
15	Trp	Prc	Asp	Val 100	Il€	His	Ser	Phe	Prc 105	Asn	Arg	Ser	Arg	Phe 11(Trp	Lys
20	His	Glu	Irp 115	Glu	Lys	His	Gly	Thr 120	ርን፡ε	Ala	Ala	Glr.	Val 125	Asp	Ala	L€u
	Asn	Ser 130	Gln	Lys	Lyrs	Tyr	Phe 135	Gly	Arg	Ser	Leu	Glu 14(Leu	Tyr	Arg	Glu
25	Leu 145	Asp	Leu	Asn	Ser	Vál 150	Leu	Leu	Lys	Leu	Gly 155	Ile	Lys	Pro	Ser	Il∈ 160
	Asn	Tyr	Tyr	Gln	Val 165	Alā	Asp	Ph€	Lys	Asp 170	Ala	L∈u	Ala	Arg	Val 175	Tyr
30	Gly	Vāl	Ile	Prc 180	Lys	IJ€	Gln	Cys	Leu 185	Pro	Prc	Ser	Gln	Asp 19(Glu	Glu
35	Val	Gln	Thr 195	Ile	Gly	Gln	Ile	G1u 200	Leu	CÀS	Leu	Thr	1ys 201	Gln	Asp	Glr.
	Gln	Leu 210		Asn	Cys	Thr	Glu 215	Pro	Gly	Glu	Gln	Pro 22(Ser	Pro	Lys	Gln
40	Glu 225	Vâl	Trp	Leu	Ala	Asn 23(Gly	Ala	Ala	Glu	Ser 23!	Arg	Gly	Leu	Arg	Val 240
	Сув	Glu	Asp	Gly	Pro 245	Val	Ph€	Tyr	Pro	Pro 250	Pro	Lys	Lys	Thr	Lys 255	His
45																
50	(2)	1NF	ORMA	NOIT.	FOR	SEÇ	ID.	NO:	181:							
			(i)	SEQU	ENCE						. ds					
55			(xi)		(B) I (D) I DUENC	'CPOI	.OGY :	lir	nea:	EÇ I	D NC): 18	31:			
	Met 1		Pro	Leu	Leu L	Leu	Gln	Leu	Ala	Val	Leu	Gly	'Ala	Ala	Leu 15	Ala
60																

	Ala	Alā	Ala	Leu 20	Val	L∈u	Ile	Ser	Ile 25	Val	Ala	Ph€	mhr	Thr .	Ala	Thr
5	Lys	Met	Prc 35	Ala	L€u	His	Arg	His 40	Glu	Glu	Glu	Lys	Phe 45	Phe	Leu	Asr.
	Ala	Lys 50	Gly	Gln	Lys	Glu	Thr 55	Leu	Prc	Ser	Ile	Trp 60	Asp	Ser	Pro	Thr
10	Lys 65	Gln	Leu	Ser	Val	Val 70	Val	Pro	Ser	Tyr	Asn 75	Glu	Glu	Lys	Arg	Leu 80
1.5	Pro	Val	Met	Met	qsA !3	Glu	Ala	Leu	Ser	Туr 90	Leu	Glu	Lys	Arg	Gln 95	Lys
15	Arg	Asp	Pro	Ala 100	Ph€	Thr	Tyr	Glu	Val 105	Ile	Val	Val	Asp	Asp 110	Gly	Ser
20	Lys	Asp	Gln 115	Thr	Ser	Lγε	Val	Ala 120	Phe	Lys	Tyr	СЭЕ	Gln 125	Lys	Tyr	Gly
	Ser	Asp 130		Vāl	Arg	Val	Ile 135	Thr	Leu	Val	Lys	Asn 14(Arg	Gly	Lys	Gly
25	Gly 1 4 5	Ala	Ile	Arg	Met	Gly 15(Ile	Phe	Ser	Ser	Arg 155	Gly	Glu	Lys	Ile	Leu 160
20	Met	Ala	Asp	Ala	Asp 165	Gly	Ala	Thr	Lys	Phe 170		Asp	Val	Glu	Lys 175	Leu
30	Glu	Lys	Gly	Leu 180	Asn	Asp	Leu	Gln	Pro 185	Trp	Pro	Asn	Gln	Met 190	Ala	Il€
35	Ala	Cys	Gly 195		Arg	Ala	Eis	Leu 200		Lys	Glu	. Ser	11e 205	Ala	Gln	Arg
	Ser	Tyr 210		Arg	Thr	Leu	Leu 215		Tyr	Gly	Phe	His 22(Phe	Leu	Val	Trp
40	Phe 225		Cys	: Val	. Lys	Gly 230		≥ Arg	Asp	Thr	Gln 235		Gly	Phe	Lys	Leu 240
45	Ph∈	Thr	Arg	, Glu	1 Ala 245	Ala	Ser	Arg	Thr	Phe 250		Ser	Leu	His	Val 255	
42	Arg	Trp) Ala	Phe 260	e Asp	Val	Glı	ı Lev	Leu 265		∶Ile	Ala	Gln	Phe 270	Phe	. Lys
50	Il∈	e Pro	275		a Glu	ı Ile	Ala	a Val 280		Trp	Thr	c Glu	1 lle 285	Glu	Gly	/ Ser
	Lys	290		l Pro	o Ph∈	e Trp	Se:) Leu	ı Glr	n Met	300 300		asp) Leu	ı Leu
55	Ph∈ 305		e Arg	g Lei	u Arg	31(u Thi	c Gly	/ Ala	a Trp 315		g Lev	ı Glu	ı Glr	320
	Arg	g Ly:	s Met	t Ası	n											

And the state of the same of

```
(2) INFORMATION FOR SED ID NO: 182:
            (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 47 amine acids
                   (B) TYPE: amino acid
                   (D) TCPCLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 182:
10
     Met Asp Ile Cys Phe Phe His Tyr Val Leu Leu Phe Phe Leu Val Arg
                            10
     Cys Ala Leu Val Val Leu Ile Leu Leu Cys Gln Gly Trp Gly Asn Gly
15
     Gly Gly Cys Val Gly Arg Val Leu Ile Ile Val Phe Ser Ser Val
                      40
20
      (2) INFORMATION FOR SEQ ID NO: 183:
            (i) SEQUENCE CHARACTERISTICS:
25
                   (A) LENGTH: 93 amino acids
                   (B) TYPE: amino acid
                   (D) TOPCLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 183:
30
     Met Ala Ser Leu Gly His Ile Leu Val Phe Cys Val Gly Leu Leu Thr
                               1(
     Met Ala Lys Ala Glu Ser Pro Lys Glu His Asp Pro Phe Thr Tyr Asr
                            25
35
     Tyr Gln Ser Leu Gln Ile Gly Gly Leu Val Ile Ala Gly Ile Leu Phe
                                40
      Ile Leu Gly Ile Leu Ile Val Leu Ser Arg Arg Cys Arg Cys Lys Phe
40
                            55
     Asn Gln Gln Gln Arg Thr Gly Glu Pro Asp Glu Glu Glu Gly Thr Phe
45
     Arg Ser Ser Ile Arg Arg Leu Ser Thr Arg Arg Arg Xaa
                     5 5
50
      (2) INFORMATION FOR SEQ ID NO: 184:
            (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 168 amino acids
                   (B) TYPE: amino acid
55
                   (D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 184:
      Met Xaa Thr Lys Glu Phe Gly Xaa Gly Arg Ala Val Gln Gln Val Leu
       16
60
```

ENCL LIVE OFFICERAL

(1) (G) + (4,4) (A)

	Asn	lie	Glu	20 20	Leu	Arg	Asp	Fne	Leu 25	Thr	Pro	PIC	ьеа	3(ser	Va.
5	Arç	Ph∈	Arg 35	Tyr	Val	Gly	Ala	Pro 4(Gln	Ala	Leu	Thr	Leu 45	Lys	Leu	Pro
	Val	Thr 5(Xāā	Asn	Lys	Ph∈	Ph∈ 55	Gln	Pro	Thr	Glu	Met 60	Ala	Ala	Gln	Ast
10	Phe 65	Phe	Gln	Arg	Trp	Lys 70	Gln	L€u	Ser	Leu	Pro 75	Gln	Gln	Glu	Ala	Glr. 80
15	Lys	Il€	Phe	Lys	Ala 85	Asn	His	Prc	Met	Asp 9(Ala	Glu	Val	Thr	Lys 9t	Alε
10	Lys	Leu	Leu	Gly 100	Ph€	Gly	Ser	Ala	Leu 10t	Leu	Asp	Asn	Val	Asp 110	Pro	Asr.
20	Pro	Glu	Asn 115	Phe	Val	Gly	Ala	Gly 120	Il€	Ile	Gln	Thr	Lys 125	Ala	Leu	Glr
	Val	Gly 13(Che	Leu	Leu	Arg	Leu 135	Glu	Frc	Asn	Ala	Gln 140	Ala	Gln	Met	Туз
25	Arg 145	Leu	Thr	Leu	Arg	Thr 150	Ser	Lys	Glu	Pro	Val 155	Ser	Arg	His	Leu	Cys 160
30	Glu	Leu	Leu	Ala	Gln 165	Gln	Phe	Xaē								
	(2)	INF	ORMA'													
35				(A) I B) I D) I	YPE :	H: 4 ami OGY:	3 am no a lin	ino cid ear	: acid EQ I		: 18	5:			
40	Met 1	Phe	Tyr	Val	Leu 5	Ser	Val	Ser	Pro	Leu 1(Leu	Xaa	Phe	Leu	Ala 15	Cys
45	Gly	Leu	Cys	Leu 20	Cys	Val	Asn	Trp	Lys 25	Ile	Ala	Ile	Ser	Gln 30	Leu	Sex
	Leu	Ser	Phe 35	Lys	Asn	Glu	Leu	Glu 4(Lys	Pro	Xaa					
50																
	(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:	186:							
55			(<u>i</u>)		(A) I (B) T	CHA LENGI TYPE: TOPOI	TH: 5	9 an Ino a	ninc cic	ació	is					
				_						EQ I						C -
60	Met	Lys	Leu	Phe	Asp	Ala	Ser	Pro	Thr	Phe	Phe	Ala	Phe	Leu	Leu	GI

	2				Ē					1(15	
_	His	Πε	Leu	Ala 20	Met	Glu	Vāl	Leu	Ala 25	Trp	Leu	Leu	Il€	Tyr 30	Leu	Leu
5	Gly	Fro	Gly 35	Trp	Vāl	Prc	Ser	Ala 4(Leu	Xaa	Arg	Leu	His 45	Prc	Gly	His
10	L€u	Ser 50	Gly	Ser	Val	Leu	Val ţ:	Ser	Ala	Ala	Ха́є					
15	(2)	INF	ORMAC	SEQU	ENCE	СНА	RACT	ERIS	IICS							
20			(xi)	(B) T D) T	YPE : OPOL	ami OGY:	no a lin	ea:			: 18	7 :			
	Met I	Asp	Vāl	Asn	Ile 5	Ala	Frc	Leu	Arg	Ala 10	Trp	Asp	Asp	Phe	Phe 1!	Pro
25	Gly	Ser	Asp	Arg 20	Phe	Ala	Arg	Prc	Asp 25	Phe	Arg	Asp	Ile	Ser 30	Lys	Trp
20	Asn	Asn	Arg 35	Val	Val	Ser	Asn	Leu 4(Leu	Tyr	Ţyr	Gln	Thr 45	Asr.	Tyr	Leu
30	Vāl	Val 5(Ala	Ala	Met	Met	Ile 55	Ser	Il€	Val	Gly	Ph∈ 60	Leu	Ser	Pro	Ph∈
35	Asn 6!	Met	Ile	Leu	Gly	Gly 70	Il∈	Val	Vāl	Val	Leu 75	Val	Phe	Thr	Gly	Phe 8(
	Val	Trp	Ala	Ala	His 85	Asn.	Lys	Asp	Val	Leu 90	Arg	Arg	Met	Lys	è: Tàs	Arç
40	Tyr	Pro	Thr	Thr 100	Phe	Val	Met	Val	Val 105	Met	Leu	Ala	Ser	Tyr 11(Phe	Leu
45	lle	Ser	Met 115	Ph€	Gly	Gly	Val	Met 120	Val	Phe	Val	Phe	Gly 12t	Ile	Thr	Ph€
-13	Pro	Leu 130	Leu	Leu	Met	Ph∈	Ile 135	His	Ala	Ser	Leu	Arg 140	Leu	Arg	Asn	Leu
50	Lys 145	Asn	Lys	Leu	Glu	Asn 150	Lys	Met	Glu	Gly	Ile 155	Gly	Leu	Lys	Arg	Th: 160
	Pro	Met	Gly	Ile	Val 165	Leu	Asp	Ala	Leu	Glu 170	Gln	Gln	Glu	Glu	Gly 175	$Il\epsilon$
55	Asn	Arg	Leu	Thr 180	Asp	Tyr	Ile	Ser	Lys 185	Val	Lys	Glu	Xaa			
60	(2)	INFO	ORMA"	NOIT	FOR	SEQ	ID 1	VC: 1	.88:							

And Control of the Control

5				C	A) L B) T D) T	ENGTI YPE : OPOLA	H: 1 ami: OGY:	46 ar no a line	mino cid ear	aci		: 188	(·			
10	M∈t 1	Ph∈	Leu	Thr	Arg £	ll∈	Leu	CÀ.E	Pro	Thr 10	Тут	Ile	Ala	Leu	Thr 15	Ph€
10	L€u	Val	Tyr	Ile 20	Val	Ala	Leu	Val	Ser 25	Gly	Gln	Leu	Cys	Met 30	Glu	Ilε
15	Ala	Arg	Gly 35	Asn	Il∈	Phe	Fhe	Leu 40	Asn	Glu	Leu	Val	Thr 45	Thr	Phe	CAE
	C7.2	Ser 50	Cys	Leu	Leu	Leu	Ser 55	Val	Pro	Tyr	Leu	His 60	Pro	Gly	Phe	Phe
20	Ψ/π €5	Ser	Ser	Leu	Cys	Lys 7(Càe	CÀE	Phe	Val	Leu 75	Val	Val	Leu	Ser	Arg 80
25	Ile	Gly	Ser	Val	Asn 85	Glu	Thr	Trp	Ser	Cys 90	Asn	Phe	Ser	Ile	Cys 95	Ser
<i>23</i>	Τγτ	Leu	lle	Ph∈ 100	Gly	Ser	Pro	Ile	Phe 105	Thr	Ala	Val	Ile	Pro 11(Lys	Arg
30	C2.s	Ala	Leu 115	Glu	Asp	Jl€	Gln	Asn 120	Asn	Pro	Ile	Gly	Cys 12:	Leu	Leu	Arg
	Cira	Thr 130	Pro	Ala	Trp	Glu	Thr 135	Glu	Gly	Asp	Ser	lle 14(Ser	Lys	Lys	Il€
35	Lys 145	Lys														
40	(2)	INF	ORMA'	TION	FOR	SEQ	ID I	NO: 3	189:							
45			(i)	(A) L B) T	CHA ENGT YPE:	H: 8 ami	4 am no a	ino cid		is					
	Met	Gly	(xi) Ser	SEQ										Phe	Ser	Phe
50	2				ŗ					10					15	
			Leu	20					25					30		
55	Glu	Ser	Gln 35	Gly	Val	Cys	Ser	Lys 40	Gln	Thr	Leu	Val	Val 4!	Pro	Leu	His
	Tyr	Asn 50	Glu	Ser	Tyr	Ser	Gln 55	Pro	Val	Tyr	Lys	Pro 60	Tyr	Leu	Thr	Leu
60	Cys	Ala	Gly	Ser	Ala	Ser	Ala	Ala	Leu	Thr	Gly	Pro	Cys	Thr	Ala	Leu

ENSTRUCT IN GOVERNMENT

	65					70					75					33
	Cys	Gly	Gly	Arç												
5																
	(2)	INFO	RMAT	ION	FOR	SEQ	ID N	10: 1	90:							
10				(1	A) L: B) T D) T	ENGTI YPE: CPOLA	H: 50 amin DGY:	8 am no ac line	ino a cid ear	acid		: 190	.			
15	Met 1	Met	Gly	Vāl	řen Ē	Gln	Leu	Leu	His	Il∈ 1(Phe	Trp	Ala	Тут	Leu 15	${\tt Il}\epsilon$
20	Leu	Arg	Met	Ala 20	His	Lys	Ph€	IJ€	Thr 25	Gly	Lys	L€u	Val	Glu 30	Asp	Glu
	Arg	Ser	Thr 35	Gly	Lys	Lys	Gln	Arg 40	Ala	Gln	Arg	Gly	Arg 45	Arg	Leu	Glr.
25	Leu	Gly 50	Glu	Glu	Gln	Arg	Ala 55	Gly	Pro	Xae						
30	(2)			rion Seçui (ENCE		RACT:	ERIS	rics		ā ؛					
35			(xi)	(B) T D) T	YPE: OPOL	ami OGY:	no a lin	cid ear			: 19	1:			
1 0	Met 1	Arg	Arq	Leu	Val į	His	Asp	Leu	Leu	Pro 1(Pro	Glu	Val	Cys	Ser 15	Leu
	Leu	Asn	Pro	Ala 20	Ala	Ile	Tyr	Ala	Asn 25	Asn	Glu	Il∈	Ser	Leu 3(Arg	Asţ
15	Val	Glu	Val 35	Tyr	Gly	Phe	Asp	Tyr 40	Asp	Tyr	Thr	Leu	Ala 45	Gln	Tyr	Alā
	Asp	Ala 50	Leu	His	Pro	Glu	Ile 55	Phe	Ser	Thr	Ala	Arg ((Asp	Ile	Leu	Iì€
50	Glu 65	His	Tyr	Lys	Tyr	Pro 70	Glu	Gly	He	Arg	Lys 7!	Tyr	Asp	Tyr	Asn	Pro 80
55	Ser	Phe	Ala	lle	Arg 38	Gly	Leu	His	Тут	Asp 90	Ile	Gln	Lys	Ser	Leu 95	Leu
	Met	Lys	lle	Asp 10(Ala	Phe	His	Tyr	Val 105	Gln	Leu	Gly	Thr	Ala 110	Tyr	Arą
60	Gly	Leu	Gln 115	Prc	Val	Fro	Asp	Glu 120	Glu	Val	Ile	Glu	Leu 125	Tyr	Gly	Gly

	Thr	Gln 130	Eis	Ile	Pro	Leu	135 135	Gln	Met	Ser	Gly	Ph∈ 14(ТУТ	Gly	Lys	G17.
5	Pro 145	Ser	∏e	Lys	Gln	Phe 150	Met	Asp	Ile	P'ne	Ser 155	Leu	Pro	Glu	Met	Ala 160
10	Leu	Leu	Ser	CAE	Val 1€5	Val	Asp	Tyr	Phe	Leu 17(Gly	His	Ser	Leu	Glu 175	Ph€
	Asp	Gln	Ala	His 180	Leu	Tyr	Lys	Asp	Val 185	Thr	Asp	Ala	Ile	Arg 190	Asp	Val
15	His	Val	Lys 195	Gly	Leu	Met	Tyr	Gln 200	Trp	Ile	Glu	Gln	Asp 205	Met	Glu	Lyf
	Tyr	11e 210	Leu	Arg	Gly	Asp	Glu 215	Thr	Phe	Ala	Val	Leu 220	Ser	Arg	Leu	Val
20	Ala 225	His	Gly	Lys	Gln	Leu 230	Phe	Leu	Il€	Thr	Asn 231	Ser	Pro	Phe	Ser	Ph∈ 240
25	Vāl	Asp	Lys	Gly	Met 245	Arg	His	Met	Val	Gly 25(Pro	Asp	Trp	Arg	His 255	Ser
	Ser	Met	Trp	Ser 260	L€u	Ser	Arg	Gln	Thr 265	Ser	Pro	Ala	Ser	Ser 270	Leu	Thr
30	Gly	Alā	Ser 275	Ph∈	Хаа	Glu	Asn	Ser 280	Met	Arg	Arg	Ala	His 285	Phe	Ser	Gly
	Thr	Gly 290	Ser	Pro	Ala	Trp	Lys 295	Arg	Alā	Arg	Ser	11e 30(Gly	Arg	Glu	Thr
35	Cys 305	Leu	Thr	Ser	Tyr	Ala 310	Xaa									
40	(2)	INF	ORMA'	NOLT	FOR	SEQ	ID I	NO: :	192 :							
45			(i)	(A) L B) T	ENGT YPE :	H: 3 ami				ā:					
			(xi)						N: S	EÇ I	D NO	: 19	2 :			
50	Met 1	Asn	Trp	Glu	Leu ţ	Leu	Leu	Trp	Leu	Leu 10	Val	Leu	Cys	Ala	Leu 15	Leu
	Leu	Leu	Leu	Val 20	Gln	Leu	Leu	Arg	Phe 25	Leu	Arg	Ala	Asp	Gly 30	Asp	Leu
55	Thr	Leu	Leu 35	Trp	Ala	Glu	Trp	Gln 40	Gly	Arg	Arg	Pro	Glu 45	Trp	Glu	Leu
	Thr	Asp 50	Met	Val	Val	Trp	Val 55	Thr	Gly	Ala	Ser	Ser 60	Gly	lle	Gly	Glu
60	Glu	Leu	Ala	Туг	Gln	Leu	Ser	Lys	Leu	Gly	Val	Ser	Leu	Val	Leu	Ser

	€ 5					70					75					80
-	Ala	Arg	Arç	Vāl	His 85	Glu	Leu	Glu	Arg	Val 9(Ĺy's	Æg	Arg	Cys	Leu 95	Glu
5	Asn	Gly	Asn.	Leu 100	Lyε	Glu	Lys	Asp	Ile 10t	Leu	Val	Leu	Pro	Leu 110	qsA	Leu
10	Thr	Asp	Thr 11!	Gly	Ser	His	Glu	Ala 120	Ala	Thr	Lys	Ala	Val 125	Leu	Gln	Glu
	Ph∈	Gly 130	Arg	Il€	Asp	Il∈	Leu 135	Val	Asn	Asn	Gly	Gly 140	Met	Ser	Gln	Arç
15	S∈r 145	Leu	CAE	Met	Asp	Thr 150	Ser	Leu	Asp	Val	Tyr 155	Arg	Lys	Leu	Il€	Glu 16(
20	Leu	Asn	Tyr	Leu	Gly 165	Thr	Val	Ser	Leu	Thr 170	Lys	Сув	Val	Leu	Pro 175	His
20	Met	ïle	Glu	Arg 180	Lys	Gln	Gly	Lys]l∈ 185	Val	Thr	Val	Asn	Ser 190	Ile	Leu
25	Gly	Il∈	11e 195	Ser	Val	Pro	Leu	Ser 200	lle	Gly	Tyr	Cys	Ala 205	Ser	Lys	His
	Ala	Leu 21(Arg	Gly	Phe	Phe	Asn 215	Gly	Leu	Arg	Thr	Glu 220	Leu	Ala	Thr	Туз
30	Pro 225	Gly	Il€	Il€	Val	Ser 230	Asn	Ile	Cys	Pro	Gly 23t	Pro	Val	Gln	Ser	Asr. 24(
35	Il€	Val	Glu	Asn	Ser 245	Leu	Ala	Gly	Glu	Val 25(Thr	Lys	Thr	Ile	Gly 255	Ast.
33	Asn	Gly	Asp	Gln 260	Ser	His	Lys	Met	Thr 261	Thr	Ser	Arg	Cys	Val 270		Leu
40	M∈t	Leu	11e 275	Ser	Met	Ala	Asn	Asp 280		Lys	Glu	Val	Trp 285	Ile	Ser	Glu
	Gln	Prc 290		Leu	Phe	Ser	Asn 295		Ph∈	Val	Ala	I1∈ 300	His	Ala	Asn	Leu
45	Gly 305		: Val	Asp	: Asn	Gln 310		Asp	Gly	Glu	Glu 315	Lys	: Asp	Xaa		
50	(2)	INF	CRMA	10 ET.	I FOR	R SEÇ) ID	NO:	193:							
55			(i)		JENCI (A) 1 (B) '	LENG TYPE	TH: ! : am	53 ar ino a	mino acic		ãs.					
JJ			(xi) SE	DUEN					SEQ :	ID: NO	o: 19	93:			
60	Met 1		Pro	Sei	r Phe		Glr	n Val	. Arg	y Val		/ Sei	r Phe	e Leu	: Phe 19	

```
Ile Leu Phe Phe Ser Phe Gly Ser Ser Ser Leu Pro Pro Gly Leu Pro
     Pro Pro Ala Ser Leu Leu Cys Cys Ala Val Gln Trp Gly Ala Arg Ala
 5
                                4(
     Leu Phe Leu Pro Ala
        5(
10
      (2) INFORMATION FOR SEQ ID NO: 194:
             (i) SEQUENCE CHARACTERISTICS:
15
                    (A) LENGTH: 42 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 194:
20
     Met Leu Val Thr Cys Ser Val Cys Cys Tyr Leu Phe Trp Leu Ile Ala
      Ile Leu Ala Gln Leu Asn Pro Leu Phe Gly Pro Gln Leu Lys Asn Glu
                  20
25
      Thr Ile Trp Tyr Leu Lys Tyr His Trp Pro
30
      (2) INFORMATION FOR SEQ ID NO: 195:
             (i) SEQUENCE CHARACTERISTICS.
                    (A) LENGTH: 102 amino acids
35
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 195:
     Met Glu Gly Thr Glu Met Gly Ala Arg Pro Gly Gly His Pro Gln Lys
40
     Trp Ser Phe Leu Trp Ser Leu Ala Leu Trp Leu Pro Leu Ala Leu Se:
45
     Val Ser Leu Phe Leu Gly Leu Ser Leu Ser Pro Pro Gln Pro Gly Leu
      Ser Leu Trp Cys Thr Leu Ser Tyr Cys Cys Glu Gln Trp Lys Phe Lys
50
      Gly Thr Pro Ser Pro Ala Leu Leu Asn Leu Gly Thr Gln Pro Lys Lys
                          70
      Asp Lys Lys Leu Glu Asp Ser Ile Ala Thr Gln Leu Arg Glu Leu Pro
55
                                  9(
     Glu Lys Asn Ser Asn Xaa
              100
```

Contact A

```
(2) INFORMATION FOR SEC ID NO: 196:
           (i) SEQUENCE CHARACTERISTICS:
5
                  (A) LENGTH: 45 amino acids
                  (E) TYPE: amino ació
                  (D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19€:
     Met Ala Leu Thr Phe Leu Leu Val Leu Leu Thr Leu Ala Thr Ser Ala
10
                                     10
     His Gly Cys Thr Glu Thr Ser Asp Ala Gly Arg Ala Ser Thr Gly Gly
15
     Fro Gln Arg Thr Ala Arg Thr Gln Trp Leu Leu Cys Xaa
20
     (2) INFORMATION FOR SEQ ID NO: 197:
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 355 amino acids
25
                  (B) TYPE: amino ació
                  (D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 197:
     Met Gly Pro Ser Thr Pro Leu Leu Ile Leu Phe Leu Leu Ser Trp Se:
30
                                      10
     Gly Pro Leu Gln Gly Gln Gln His His Leu Val Glu Tyr Met Glu Arg
               20 25 30
     Arg Leu Ala Ala Leu Glu Glu Arg Leu Ala Gln Cys Gln Asp Gln Se:
35
     Ser Arg His Ala Ala Glu Leu Arg Asp Phe Lys Ash Lys Met Leu Pro
                        5 5
40
     Leu Leu Glu Val Ala Glu Lys Glu Arg Glu Ala Leu Arg Thr Glu Ala
                 70
     Asp Thr Ile Ser Gly Arg Val Asp Arg Leu Glu Arg Glu Val Asp Tyr
45
          85 90
     Leu Glu Thr Gln Asn Pro Ala Leu Pro Cys Val Glu Phe Asp Glu Lys
     Val Thr Gly Gly Pro Gly Thr Lys Gly Lys Gly Arg Arg Asn Glu Lys
50
                   12(
      Tyr Asp Met Val Thr Asp Cys Gly Tyr Thr Ile Ser Gln Val Arg Ser
                    135
55
      Met lys Ile Leu Lys Arg Phe Gly Gly Pro Ala Gly Leu Trp Thr Lys
                                155
              150
      Asp Fro Leu Gly Gln Thr Glu Lys Ile Tyr Val Leu Asp Gly Thr Gln
60
                                    170
                    165
```

Contract Contract Contract

	Asn	Asp	Thr	Ala 180	Ph€	Vāl	Ph€	Pro	Arg 185	Leu	Arg	Asp	Phe	Thr 19(Leu	Alā
5	Met	Alā	Ala 195	Arg	Lys	Ala	Ser	Arg 200	Val	Arg	Val	Prc	Phe 201	Pro	Trp	Val
10	Gly.	Thr 210	Gly	Gln	L€u	Val	ገንጥ 215	Gly	Gly	Phe	Leu	Tyr 22(Phe	Ala	Arg	Arg
10	Pro 225	Pro	Gly	Arg	Pro	Gly 23(Gly	Gly	Gly	Glu	Met 235	Glu	Asn	Thr	Leu	Gln 240
15	Leu	Il€	Lys	Phe	His 24:	Leu	Ala	Asn	Arg	Thr 250	Val	Val	Asp	Ser	Ser 255	Val
	Ph€	Pro	Ala	Glu 260	Gly	Leu	lle	Pro	Pro 265	Tyr	Gly	Leu	Thr	Ala 270	Asp	Thr
20	Tyr	lle	Asp 275	Leu	Ala	Ala	Asp	Glu 280	Glu	Gly	Leu	Trp	Ala 285	Val	Tyr	Ala
25	Thr	Arg 290	Glu	Asp	Asp	Arg	His 295	Leu	Cys	Leu	Ala	1ys	Leu	Asp	Pro	Gln
	Thr 30t	L∈u	Asp	Thr	Glu	Gln 31(Gln	Trp	qaA	Thr	Pro 315	Cys	Pro	Arg	Glu	Asn 320
30	Ala	Glu	Ala	Ala	Phe 325	Val	Ile	Cys	Gly	Thr 330	Leu	Тут	Val	Val	Tyr 335	Asn
	Thr	Arg	Pro	Ala 340	Ser	Arg	Ala	Arg	Ile 345	Gln	Cys	Ser	Phe	Asp 35(Ala	Se:
35	Gly	Pro	Xaa 355													
40	(2)	INF				SEQ										
45			(i)	•	(A) I (B) T	CHA LENGT TYPE: TOPOL	TH: 7	74 an Inc a	nino acid		ls					
4.			(xi)			E DE				EQ I	D NC): 19	31			
50	Met 1		Leu	Pro	Leu <u>t</u>	Leu	Il∈	Phe	Val	Leu 10		Pro	Lys	Val	Val 15	Asn
	Thr	Ser	Asp	Pro 20		Met	Arg	Arg	Glu 25		Glu	Gln	Ser	Met 30		Met
55	Leu	Asn	Ser 35		His	Glu	Leu	Pro 40		: Val	Ser	Glu	Phe 4!	Met	Thr	Arg
	Leu	Ph∈ 50		Ser	Lys	Ser	Ser 55		Lys	Ser	Ser	Ser 60		ser '	Ser	Lys
60	Thr	Gly	. Lys	Ser	Gly	Ala	Gly	Lys	Arg	Arç	:					

7(€Ē (2) INFORMATION FOR SEQ ID NO: 199: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 113 aminc acids (E) TYPE: amino acid 10 (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 199: Met Phe Thr Met Leu Cys Ile Ash Gly Thr Thr Pro Arg Pro Leu Pro 1(15 Val Pro Ser Fro Phe Gly Cys Met Ile Phe Phe Phe Lys Asn Pro 25 Trp Lys Gln Arg Leu Leu Gln Gly Trp Leu Gly Ala Arg Prc Ile His 20 40 Leu Leu Gly Tyr Leu Pro Leu Ser Leu Leu Trp Cys Pro Phe Pro Leu 25 Pro Cys Ala Arg Cys Ser Val Val Tyr Ile Ser Ser Pro Arg His Gly Ala His Ala Frc Arg Asp Met Ile Leu Ser Leu Val Leu Ala His Gly 30 Ala Leu Tyr Lys Glu Leu Gly Gly Arg Gly Arg Lys Trp Glu Pro Ser 105 Xaa 35 (2) INFORMATION FOR SEQ ID NO: 200: 40 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 123 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 200: Met Ala Cys Arg Cys Leu Ser Phe Leu Leu Met Gly Thr Phe Leu Ser 50 Val Ser Gln Thr Val Leu Ala Gln Leu Asp Ala Leu Leu Val Phe Pro Gly Gln Val Ala Gln Leu Ser Cys Thr Leu Ser Pro Gln His Val Thi 55

Ile Arg Asp Tyr Gly Val Ser Trp Tyr Gln Gln Arg Ala Gly Ser Ala

Pro Arg Tyr Leu Leu Tyr Tyr Arg Ser Glu Glu Asp His His Arg Pro

55

7(

42.00

	Ala	Asp	Il€	Frc	Asp Et	Arg	Phe	Ser	Ala	Ala 9(Lys	Asp	Glu	Ala	His 95	Asr.
5	Ala	ርን.ε	Val	Leu 100	Thr	Il∈	Ser	Pro	Val 105	Gln	Pro	Glu	Asp	Asp 110	Ala	Asr
10	Tyr	туr	Cys 115	Ser	Val	Gly	Tyr	Gly 120	Phe	Ser	Pro					
	(2)	INFO	TAMAC	NOI	FOR	SEQ	ID 1	NO: 2	201 :							
15 20			(i) s	(. ()	A) L: B) T D) T	ENGT: YPE : OPOL	H: 3 ami OGY:	15 a no a lin	mino cid ear	aci		: 201	1:			
20	Met 1	Ala	Gly	Gly	Arg t	Сув	Gly	Pro	Xaa	Leu 1(Thr	A.ē	Leu	L€u	Ala 15	Ala
25	Trp	Ile	Ala	Ala 20	Val	Ala	Ala	Thr	Ala 25	Gly	Pro	Glu	Glu	Ala 3C	Ala	Leu
	Prc	Pro	Glu 35	Glr.	Ser	Αrg	Val	Gln 40	Pro	Met	Thr	Ala	Ser 4	Asn	Trp	Thi
30	Leu	Val 50	Met	Glu	Gly	Glu	Trp 55	Met	Leu	Lys	Phe	Tyr 6(Ala	Pro	Trp	CAE
35	Prc 65	Ser	CAE	Gln	Gìr	Thr 70	Asp	Ser	Glu	Trp	Glu 7!	Ala	Ph€	Ala	Lys	Asn 80
55	Gly	Glu	lle	Leu	Glr. E!	Ile	Ser	Val	Gly	Lys 90	Val	Asp	Val	Ile	Gln 95	Glu
40	Pro	Gly	Leu	Ser 100	Gly	Arg	Phe	Phe	Val 105	Thr	Thr	Leu	Pro	Ala 110	Phe	Ph€
	His	Ala	Lys 115	Asp	Gly	lle	Phe	Arg 120	Arg	Tyr	Arg	Gly	Pro 125	Gly	Ile	Ph€
45	Glu	Asp 130		Gln	Asn	Tyr	Ile 135		Glu	Lys	Lys	Trp 14(Gln	Ser	Val	Glu
50	Pro 145		Thr	Gly	Trp	Lys 150		Pro	Ala	Ser	Leu 15:	Thr	Met	Ser	Gly	Met 160
30	Ala	Gly	Leu	Phe	Ser 165	Ile	Ser	Gly	Lys	Ile 17(His	Leu	His	Asn 175	туг
55	Phe	Thr	· Val	Thr 180	Leu	Gly	·Il€	Pro	Ala 185	Trp	Cys	Ser	Tyr	Val 190		Phe
	Val	Il€	Ala 195	Thr	L€u	. Val	Ph€	Gly 200		Ph∈	Met	Gly	Leu 205		Leu	Val
60	Val	ıl€	e Ser	Glu	Cys	Ph∈	Tyr	. Val	Pro	Leu	Pro	Arg	His	Lev	Ser	Glu

		21(215					220				
5	Arg 225	Ser	Glu	Gln	Asr.	Arg 230	Arg	Ser	Glu	Glu	Ala 235	His	Arg	Ala	Glu	Gl:. 24(
J	Leu	Glr.	Yet	Ala	Glu 245	Glu	Glu	Lys	Asp	Asp 25(Ser	Asr.	Glu	Glu	Glu 25!	ASI.
10	Lys	Yer	Ser	Leu 260	Val	Asp	Asr	Glu	Glu 2€5	Glu	Lys	Glu	Asp	Leu 270	Gly	As;
	Glu	Asp	Glu 27 <u>1</u>	Ala	Glu	Glu	Glu	Glu 280	Glu	Glu	Asp	Asn	Leu 285	Ala	Ala	Gly
15	Val	Asp 29(Glu	Glu	Arg	Ser	Glu 295	Ala	Asr.	Asp	Gln	Gly 300	Pro	Pro	Gly	Glu
20	Asr 305	Gly	Val	Thr	Arg	Glu 31(Xaa	Ser	Arg	Alā	Xaa 315					
25	(2)	INF	ORMAI							2						
			(xi)	(A) L B) T D) T	ENGT YPE: OPOL	H: 2 ami OGY:	36 a ne a lin	mino ció ea:	aci		. 20	ç .			
30	Met 1	Gly												Pro	Gln 15	Ea:
35	Thr	His	Il∈	Asp 2(Val	His	He	His	Gln 25	Glu	Ser	Ala	Leu	Ala 30	Lys	Let
	Leu	Leu	Thr 3:	САЕ	Cys	Ser	Ala	Leu 4(Arg	Prc	Arg	Alā	Thr 45	Gln	Alā	Αţ¢
40	Gly	Ser 5(Ser	Arg	Leu	Leu	Val 55	Ala	Ser	Trp	Val	Met 6(Gln	Ile	Val	Leu
45	Gly 6t	Ile	Leu	Ser	Ala	Val 70	Leu	Gly	Gly	Fhe	Phe 75	Tyr	Ile	Arg	Asp	8(Ty:
	Thr	Leu	L€u	Val	Thr 85		Gly	Ala	Ala	. Ile 9(Trp	Thr	Gly	Ala	Val 95	Ala
50	Val	Leu	. Ala	Gly 100	Ala	Ala	Ala	Phe	11∈ 105	: Tyr	Glu	Lys	Arg	Gly 11(· Gly	Thi
	Tyr	Trp	Ala 115	Leu	. Leu	. Arg	Thr	120		ı Ala	Leu	Ala	125		: Ser	Thi
55	Ala	11∈ 130		Ala	Leu	Lys	135) Asn	n Glu	Asp	Phe 140		Tyr	Gly	Ty
60	Ser 145	-	Tyr	Asn	Ser	Ala 150		Arg	: Ile	e Ser	Ser 155		Ser	Asp	Trp	16

Lite Garage

	Thr	Fre	Ala	Frc	Thr 165	Gln	Ser	Fre	Glu	Glu 170	Val	Arg	Arg	L€u	His 175	Leu
5	Cys	Thr	Ser	Phe 180	Met	Asr	Met	Leu	Lys 181	Ala	Leu	Phe	Arg	Thr 19(Leu	Gln
	Ala	Met	Leu 19t	Leu	Gly	Val	Trp	Il∈ 200	Leu	Leu	Leu	Leu	Ala 205	Ser	Leu	Alē
10	Frc	Leu 21(qrp	Leu	Tyr	Сує	1rp 215	Arg	M∈t	Phe	Prc	Thr 220	Lys	Gly	Lys	Arg
15	Asp 225	Gln	Lys	Glu	Met	Leu 230	Glu	Val	Ser	Gly	11e 235	Xaa				
20	(2)	INFO		SEQU () (ENCE A) L B) T D) T	CHA ENGT YPE: OPOL	RACT H: 9 ami OGY:	ERIS 3 am no a lin	TICS inc cic ear	acid			-			
25	Met	Ilε	(xi) His		UENC Gly E									Pro	Val	Ala
30		Æla	Gln	Thr 20		Pro	Gly	Glu	Arg 25		Ser	Leu	Pro	Ala 3(Ωλι
	Frc	Gly	Thr 35	Ser	Gly	Ser	CAE	Ser 4(Gly	Cys	Gly	Ser	Leu 45	Ser	Leu	Pro
35	Leu	Leu 5(Ala	Gly	Leu	Val	Ala 55	Ala	Asp	Ala	Val	Ala 60	Ser	Leu	Leu	Il€
40	Val €:	Gly	Ala	Val	Ph∈	Leu 70	Cys	Ala	Arg	Pro	Arg 75	Arg	Ser	Pro	Ala	Glr. 80
	Glu	Asp	Gly	Lys	Val 85	Tyr	Ile	Asn	Met	Pro 9(Glγ	Arg	Gly			
45	(2)	INF	ORMA'	TION	FOR	SEQ	IDI	N O: 3	204:							
50				(ENCE A) I B) I D) I UENC	ENGT YPE : OPOL	H: 3 ami OGY:	5 am no a lin	ino cid ear	acid		: 20	4 :			
55	Met :	Trp	Ser	Ala	Gly 5	Arg	Gly	Gly	Ala	Ala 10	Trp	Pro	Val	Leu	Leu 1	GJ?
	Leu	Leu	Leu	Ala 20		Leu	Val	Pro	Gly 2!	Gly	Gly	Ala	Ala	Lys 30	Thr	Gly
60	Δla	Aen	Ser													

 $\{(i,0),(i,1),(M_{i})\}\subseteq\{(i,0)\}$

3:

```
(2) INFORMATION FOR SEC ID NO: 20E:
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 43 amino acids
                   (E) TYPE: amino acid
](:
                  (D) TOPCLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 205
     Asp Cys Xaa His Val Ser Val Leu Gln Ser Thr Ile Ser Pro Leu Leu
                                     10
15
     Fro Leu Fro Leu Leu Fro His Gly Asn Cys Glu Glu Ala Pro Trp.
     Glr. Ala Ala Val Ile Gly Gly Gly Asp Arg Ile
20
             3.5
     (2) INFORMATION FOR SEQ ID NO: 206:
25
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 85 amino acids
                   (B) TYPE: amine acid
                   (D) TOPOLOGY: linear
30
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20t:
     Met Arg Asp Cys Leu Ser Leu Lys Pro Arg Pro Leu Phe Pro Thr Glm
                            10
              35
      Phe Phe Phe Ile Leu Leu Ile Phe Ile Ala Glu Val Ala Ala Ala
                 20
     Val Val Ala Leu Val Tyr Thr Thr Met Val Arg His Trp Asp Gly Gly
                      4(
40
      Arg Glu Glu Asp Trp Ala Lys Prc Trp Glu Trp Ala Val Ala Cys Glu
      Trp Pro Pro Ser Val Pro Ala Pro Lys His Trp Pro Ala Ser Pro Arg
45
      65 7(
      Leu Ser Thr Ser Xaa
50
      (2) INFORMATION FOR SEQ II NO: 207:
             (i) SEQUENCE CHARACTERISTICS:
55
                   (A) LENGTH: 208 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 207:
60
      Met His Gly Asn Glu Ala Leu Gly Arg Glu Leu Leu Leu Leu Met
```

	1				ŗ					1(15	
5	Glr	Ph€	L∈u	Cys 20	His	Glu	Fh∈	Leu	Arg 25	Хаа	Asrı	Pro	Αrg	Väl 30	Thr	Arg
•	Leu	Leu	562 35	Glu	M∈t	ΜĠ	Il€	His 40	Leu	Leu	Pro	Ser	Met 45	Asn	Prc	Asp
10	GJ7.	777 50	Glu	Π€	Ala	Tyr	H15 55	Arg	Gly	Ser	Glu	Leu €(Vāl	Gly	Trp	Ala
	Glu 65	Gly	Μā	Trp	Asr.	Asn 70	Gln	Ser	Ile	Asp	Leu 71	Asn	His	Asn	Phe	Ala 80
15	Хаа	Leu	Asn	Thr	Prc 85	Leu	Trp	Glu	Ala	Glr. 9(Asp	Asp	Gly	Lys	Val 95	Pro
20	His	∏€	Vāl	Frc 100	Asn	His	His	Leu	Pro 105	Leu	Pro	Thr	īyr	Tyr 110	Thr	Leu
20	Frc	Asn	Ala 115	Thr	Val	Ala	Pro	Glu 120	Thr	Arg	Ala	Vāl	Il∈ 125	Lys	Trp	Met
25	Lys	Arg 130	Il€	Frc	Phe	Vāl	Leu 135	Ser	Ala	Asn	Leu	His 14(Gly	Gly	Glu	Leu
	Val 145	Väl	Ser	Тут	Frc	Phe 15(Asp	Met	Thr	Arg	Thr 155	Prc	lib	Ala	Ala	Arg 160
30	Glu	Leu	Thr	Pro	Thr 165	Fro	Asp	Asp	Ala	Val 170	Ph€	Arg	Trp	Leu	Ser 175	Thi
2.5	Val	Tyr	Ala	Gly 180	Ser	Asr	Leu	Ala	Met 185	Gln	Asp	Thr	Ser	Arg 190	Arg	Pro
35	Cys	His	Ser 195	Gln	Asp	Phe	Ser	Val 200	His	Gly	Asn	∏∈	Ile 201	Asn	Gly	Ala
40																
45	(2)	INF	ORMA' (SEQ CHA				:						
				(E) T	ENGT YPE: CFCL	ami	no a	cid	acid	E					
50				SEÇ	UENC	E DE	SCRI	PTIO	N: S							
	Met 1	Glu	Il∈	Ser	Cys 5	Leu	Leu	Leu	Leu	Ile 1(Gìn	Asp	Ser	Asp	Glu 15	Met
55	Glu	Asp	Gly	Pro 2(Gly	Vāl	Gln	Asp								
60	(2)	INF	ORMA'	TION	FOR	SEÇ	ID :	NO:	209:							

5			(i) 9 (xi)	(; (;	A) Li B) T D) T	ENGTI YPE : OPCLA	i: 4 am.i: DGY:	83 ar nc a line	mino cid ea:	aci		: 209	9 :			
10	Met 1	Ala	Thr	Gly	Gly E	Gly	Il∈	Arc	Alē	Met 1	Thr	Ser	Leu	Tyr	Gly 15	Glr.
10	Leu	<i>L</i> la	Gly	Leu 20	Lys	Glu	Leu	Gly	Leu 25	Leu	Asp	Cys	Xaa	Ser 30	Tyr	Il€
15	Thr	Gly	Ala 35	Ser	Gly	Ser	Thr	Trp 40	Ala	Leu	Ala	Asn	Leu 45	Tyr	Lys	Asj.
	Prc	G] u 5(Trp	Ser	Gln.	Lys	Asp 55	Leu	Ala	Gly	Pro	Thr 60	Glu	Leu	Leu	Lys
20	Thr 65	Gln	Val	Thr	Lys	Asn 70	Lys	Leu	Gly	Val	Leu 75	Ala	Pro	Ser	Gln	Let 80
25	Gln	Arg	Tyr	Μģ	Gln 85	Glu	Leu	Ala	Glu	Arg 9(Ala	Arg	Leu	Gly	Tyr 9!	Pro
4 3	Ser	CÀE	Phe	Thr 10(Asn	Leu	Trp	Ala	Leu 10!	Ile	Asn	Glu	Ala	Leu 110	Leu	His
30	qzA	Glu	Pro 115	His	Asp	His	Lys	Leu 120	Ser	Asp	Gln	Arg	Glu 125	Ala	Leu	Sۓ
	His	Gly 13(Gln	Asn	Pro	Leu	Pro 135	Il€	Tyr	Cys	Ala	Leu 14(Asn	Thr	Lys	Gly
35	Gln 145	Ser	Leu	Thr	Thr	Phe 15(Glu	Phe	Gly	Glu	Trp 15!	Сув	Glu	Phe	Ser	Pro 160
40	Tyr	Glu	Val	Gly	Phe 16t	Pro	Lys	Tyr	Gly	Ala 17(Ph∈	Ile	Prc	Ser	Glu 175	Let
40	Ph∈	Gly	Ser	Glu 18(Phe	Phe	Met	Gly	Gln 185	Leu	Met	Lys	Arg	Leu 190	Pro	Glu
45	Ser	Arg	11€ 19!	Cys	Fh∈	Leu	Glu	Gly 200	Il∈	Trp	Ser	Asn	Leu 205	Tyr	Ala	Ala
	Asn	leu 21(Gln	Asp	Ser	Leu	Tyr 215	Trp	Ala	Ser	Glu	Pro 220		Gln	Phe	Tr
50	Asp 225	Arg	l Trp	Vāl	Arg	Asn 230		Alā	Asn	Leu	Asp 235	Lys	Glu	Gln	Val	Pr (
<i>E</i>	Leu	Leu	l Lys	lle	Glu 245	Glu	Pro	Pro	Ser	Thr 25(Ala	Gly	Arg	Ile	Ala 255	Gl [.]
55	Ph∈	Phe	Thr	Asp 260		Leu	Thr	Trp	Arg 265	Pro	Leu	Ala	Glr.	Ala 270	Thr	Hi
60	Asn	Ph€	Leu 274	Arg	Gly	Leu	His	Phe		Lys	Asp	Туг	Phe 285		His	Þr

	Eis	Ph∈ 29(Ser	Thr	Trp	Lys	Ala 295	Thr	Thr	Leu	Asp	Gly 300	Leu	Pro	Asn	Glr
5	Leu 305	Thr	Pro	Ser	Glu	Pro 310	His	Leu	Cys	Leu	Leu 315	Asp	Val	Gly	Tyr	Leu 320
10	Il€	Asr.	Thr	Ser	Cys 321	Leu	Frc	Leu	Leu	Gln 330	Frc	Thr	Arg	Asp	Val 335	As;
10	Leu	Il€	Leu	Ser 340	Leu	Asp	Tyr	Asn	Leu 345	His	Gly	Alá	Ph∈	Gln 350	Gln	Leu
15	Gln	Leu	Leu 351	Gly	Arg	Ph€	Cys	Gln 360	Glu	Gln	Gly	Ile	Pro 365	Phe	Prc	Pro
	lle	Ser 370	Frc	Ser	Pro	Glu	Glu 375	Gln	Leu	Gln	Pro	Arg 380	Glu	Суѕ	His	The
20	Ph∈ 385	Ser	Asp	Prc	Thr	Cys	Prc	Gly	Ala	Pro	Ala 395	Val	Leu	His	Ph€	Pr (
0.5	Leu	Val	Ser	Asp	Ser 405	Ph∈	Arg	Glu	īyx	Ser 41(Ala	Pro	Gly	Val	Arg 415	Arç
25	Thr	Pro	Glu	Glu 420	Ala	Ala	Ala	Gly	Glu 4 2!	Val	Asn	Leu	Ser	Ser 430	Ser	Asr
30	Ser	Pro	Tyr 435	His	Тут	Thr	Lys	Val 440	Thr	Tyr	Ser	Gln	Glu 445	Asp	Val	Asr
	Lys	Leu 45(Leu	His	Leu	Thr	His 455	Tyr	Asn	Val	Суя	Asn 460	Asn	Gln	Glu	Gl1.
35	Leu 465	Leu	Glu	Ala	Leu	Arg 4 70	Gln	Ala	Val	Gln	Arg 475	Arg	Arg	Gln	Arg	Arç 48(
40	Pro	His	Xaa													
40																
	(2)	INF	ORMA	TION	FOR	SEÇ	ID I	NO: 2	210:							
45			(i)	(ENCE A) L B) T	ENGT YPE :	H: 1 ami	3 am no a	inc cid		<u> </u>					
			(3ri)		D) T UENC					EO T	D NO	. 21	n -			
50																
	Leu I	Glu	Val	Gly	Cys 5	Il∈	Gln	Val	Ala	Pro 1(Asp	Thr	Phe			
55	(2)	INF	orma'	TION	FOR	SEÇ	ID	NO: I	211:							
			(i)	SEOU	ENCE	CHA	RACT	ERIS	TICS	:						
			/	-	A) L						٤					
60				(в) Т	YPE:	ami	no a	cic							

```
(D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEÇ ID NO: 211:
     Met Ser Led Phe Phe Led Led Thr Led The Ser Lys Led His Gly Asp
5
      Ala Glu Val Cyr
10
     (2) INFORMATION FOR SEQ ID NO: 212:
             (i) SEQUENCE CHARACTERISTICS:
15
                    (A) LENGTH: 55 amino acid:
                    (E) TYPE: aminc acid
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 212:
      Met Fro His Fro Pro Leu Pro Glu Thr Ser Leu Glu Ala Gln Leu Pro
20
      Met Gly Leu Leu Gln Leu Leu Arg Cys Ser Val Gln Ala Trp Ser Pro
                  20
                              21
25
      Pro Pro Ser Ser Phe Cys Pro Gly Ser Glu Pro Arg Ser Ala Ser Ala
                          4(
      His Trp Gly Tyr Trp Trp Pro
30
       5(
      (2) INFORMATION FOR SEC ID NO: 213:
35
             (i) SEQUENCE CHARACTERISTICS
                    (A) LENGTH: 35 amino acids
                    (B) TYPE: amino acic
                    (D) TOPOLOGY: linea:
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 213:
40
      Asp Pro Glu Thr Arg Trp His His Gly Gly Ser Ala Gln Ash Gly Lec
      Leu Met Leu Ile Ser Val Leu Gin Gin Pro Val Ile Gly Thr Gly Ser
45
      Tyr Leu Cyr
50
       (2) INFORMATION FOR SEQ ID NO: 214
 55
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 230 amino acids
                    (B) TYPE: amino acid
                    (D) TOPCLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 214:
 60
```

Appended N. 11 1991

Commence of the second

	Me t	Glu	Prc	Leu	Arg E	Leu	L∈u	lì€	Leu	Leu 10	Phe	Val	Thr	Glu	Leu 15	Ser
5	Gly	Alā	His	Asn 20	Thr	Thr	Val	Ph€	Gln 2ŧ	Gly	Val	Ala	Gly	Gln 3(Ser	Leu
	Gln	Val	Ser 35	Cys	Prc	làx	Asr	5∈r 4(Met	Lys	His	Trp	Gly 4:	Arç	Arg	Lys
10	Ala	Trp 50	Сує	Arg	Gln	Leu	êi GJA	Glu	Lys	Gly	Pro	Cys	Gln	Arg	Val	Va]
15	Ser 65	Thr	His	Asn	Leu	Trp 70	Leu	Leu	Ser	Phe	Leu 75	Arg	Arg	Trp	Asn	613.
12	Ser	Thr	Ala	Ile	Thr 85	Asp	Asp	Thr	Leu	Gly 90	Gly	Thr	Leu	Thr	Ile 95	Thi
20	Leu	Arg	Asn	Leu 100	Gln	Frc	His	Asp	Ala 105	Gly	Leu	туr	Gln	Cys 11(Gln	Set
	Leu	His	Gly 115	Ser	Glu	Ala	Asp	Thr 120	Leu	Arg	Lys	Val	Leu 125	Val	Glu	Val
25	Leu	Ala 130	Asp	Pro	Leu	Asp	His 135	Arg	Asp	Ala	Gly	Asp 14(Leu	Trp	Phe	Prc
30	Gly 145	Glu	Ser	Glu	Ser	Phe 15(Glu	Asp	Ala	His	Val 155	Glu	His	Ser	Iî∈	Ser 160
30	Arg	Ser	Leu	Leu	Glu 165	Gly	Glu	ll€	Pro	Phe 170		Pro	Thr	Ser	11e 17:	Leu
35	Leu	Leu	Leu	Ala 180		Il€	Ph€	Leu	Ile 185		: Ile	Leu	Ala	Ala 190	Ser	Xaa
	Leu	Trp	Ala 195		Ala	Trp	Eis	Gly 200	Gln	Lys	Pro	Gly	Thr 201	His	Pro	Prc
40	Ser	Glu 210		Asp	Cys	Gly	His 215		Pro	Gly	/ Tyr	Gln 220		ı Gln	Thr	Leu
45	Pro 22:	Gly	· Leu	Arg	n Asp	Th: 23(
	(2)	INF			1 FOF											
50			(i)	SEQ	(B)	E CHA LENG TYPE TOPO	TH: : am	231 ino	amin aciā		iās					
55	Met	t Glu			ouen . Arg						u Phe			r Gli	ı Lei	u Ser
60	Gly	y Ala	a His	s Ası 2		r Thi	r Val	l Phe	e Gli 2!	n G1		l Ala	a Gly	y Gl: 3:	n Se:	r Lei

	Glr.	val	Ser	CVS	Fre	Tryr	Asr	Ser	Met	l.N.S	Has	Trn	Giv	Arc	Arg	INS
	-		35	-2-		-3-		40		3-			4!		:	, -
5	Ala	Trp 50	CÀŧ	ΑΥÇ	Glr.	L∈u	Gly E!	Glu	Lys	Gly	Frc	€(Cys	Gln	Arg	Val	Val
10	er Eer	Thr	His	Asr.	leu	Ir. 70	Leu	Leu	Ser	Ph∈	Leu 7:	Arç	Arg	Trp	Asn	Gly 80
•	Ser	Thr	Ala	Il€	rufT !3	Yeb	Asp	Thr	Leu	Gly 90	Gly	Thr	Leu	Thr	11∈ 95	Thr
15	Leu	Arç	Asr.	Leu 10(Gln	Frc	His	qzA	Ala 105	Gly	Leu	Tyr	Gln	Cys 11(Gln	Se:
	Leu	His	Gly 115	Ser	Glu	Ala	Asp	Thr 120	Leu	Arg	Lys	Val	Ьеи 125	Val	Glu	Val
20	Leu	Ala 130	Asp	Frc	Leu	Asp	His 135	Arg	Asp	Ala	Gly	Asp 14(Leu	Trp	Phe	Pro
25	Gly 145	Glu	Ser	G] u	Ser	Phe 15(Glu	qsA	Ala	His	Val	Glu	His	Ser	lle	Ser 160
	Arg	Ser	Leu	Leu	Glu 1€:	Gly	Glu	Ίlε	Pro	Phe 170	Pro	Pro	Thr	Ser	Ile 175	Leu
30	L€u	Ъeu	Leu	Ala 18(Сує	ll€	P'n€	Leu	Il∈ 185	Lys	Il∈	Leu	Alā	Ala 190	Ser	Ala
	Leu	Trp	Ala 195	Ala	Ala	Trp	His	Gly 200	Gln	Lys	Pro	Gly	Thr 20:	His	Pro	Prc
35	Ser	Glu 210	Leu	Asp	Cys	Gly	His 215	Asp	Pro	Gly	Tyr	Gln 22(Leu	Gln	Thr	Leu
40	Pro 225	Gly	Leu	Arg	Asp	Thr 23(Xāč									
	(2)	INF	ORMA'	TION	FOR	SEQ	ID	NO: 1	216:							
45			(i)	(A) L B) T	ENGT YPE:	H: 1 ami	ERIS' 27 a no a lin	mino cid		āε					
50			(xi)	SEÇ	UENC	E DE	SCRI	PTIO	N: S	EQ I	D NO	: 21	t :			
50	Met :	Gly	Leu	Thr	Gly	Phe	Gly	Val	Phe	Phe 10	Leu	Phe	Ph∈	Gly	Met 15	${\tt Il}\epsilon$
55	Leu	Phe	Phe	Asp 20	Lys	Ala	Leu	Leu	Ala 25	Ile	Gly	Asn	Val	Leu 30	Phe	Val
	Ala	Gly	Leu 35	Ala	Ph∈	Val	Ile	Gly 40	Leu	Glu	Arg	Thr	Phe 4:	Arg	Phe	Ph∈
60	Phe	Gln	Lys	His	Lys	Met	Lys	Ala	Thr	Gly	Phe	Phe	Leu	Gly	Gly	Val

	50					Ēŗ					€(
,	Fh∈ Val €5	Val	L€u	ïl€	G13.	Trr	Frc	Leu	Ile	Gly 75	M∈t	Ile	Phe	Glu	11€ 80
5	Tyr Gly	Phe	Ph€	Leu E'	L€u	Phe	Arg	Gly	Ph∈ 90	Phe	Pro	Val	Val	Val șŧ	GlΣ.
10	The Ile	Ārģ	Arg 100	Val	Fic	Val	Leu	Gly 105	Ser	Leu	Leu	Asn	Leu 11(Pro	Glу
	lle Arg	Ser 115	Ph€	Val	Asp	Lys	Val 120	Gly	Glu	Ser	Asn	Asn 121	Met	Val	
15	(2) INFO	ORMAG	NOI	FOR	SEO	ID:	NO:	217:							
20		(i) :	EÇU)))	ENCE A) I B) I	CHA LENGI TYPE :	FACT TH: 4 ami	ERIS 17 am .no a	TICS nino ncid	acio		o: 21	.5 :			
25	Met Ile 1	Arg	Lys	Leu !	Eis	Lys	Ile	Ile	Val		Ser	Pro	Arg	Val 15	Il€
20	Val Leu	Leu	Asn 20		Ph€	Phe	Phe	: 11€ 25	Lys	Ala	lys	Phe	Val	Leu	. īyi:
30	Ile Phe	Val	Ph∈	His	Val	Lev	Asp 4(Ser	· Ile	e Ser	Tyr 45	Pro	Val	
35	(2) INF	ORMA	TION	: FOF	R SEÇ) ID	NO:	218:							
40				(A) (B) (D)	LENG TYPE TOPO	TH: : am LOGY	41 a ino : 1i		aci		0: 2	18.			
45	Met Lei 1	ı Lev	ı Ası	n Gli		s Pho	e Ly:	s Ile	Pho 1		y Se:	r Lei	ı Ile	⊕ His 15	s Met
	Asn Lei	ı Lev	Phe 20		a Lei	u Il	e Se:	r Let 25		y Se	r Se	r As:	n Le	u Sei	r Gly
50	Val Glr	n Phe		s Cy	s Gl	u Th	r Va 4		a						
55	(2) IN	FORM	ATIO	n fo	R SE	Q ID	NO:	219	:						
		(i)	SEÇ	(A)	LENG	GTH:	105	STIC amin acid	o ac	cids					
60								inear							

```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 219:
     Met Glm Fro Leu Ash Phe Ser Ser Thr Xaa Cys Ser Ser Phe Ser Fro
 5
      Fro Thr Thr Val Ile Leu Leu Ile Leu Leu Cys Fne Glu Gly Leu Leu
      Phe Leu Ile Phe Thr Ser Val Met Phe Gly Thr Gln Val His Ser Ile
10
      Cys Thr Asp Glu Thr Gly Ile Glu Gln Leu Lys Lys Glu Glu Arg Arg
                            E, E
15
     Trp Ala Lys Lys Thr Lys Trp Met Asn Met Lys Ala Val Phe Gly His
     Fro Fhe Ser Leu Gly Trp Ala Ser Pro Phe Ala Thr Pro Asp Gln Gly
20
     Lys Ala Asp Pro Tyr Gln Tyr Val Val
               2 € (
25
      (2) INFORMATION FOR SEQ ID NO: 220:
            (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 29 amino acid:
30
                   (P) TYPE: amino acid
                   (D) TCPOLOGY: linea:
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 220:
     Met Tyr Thr Asn His Phe Asn Leu Tyr Leu Lys Tyr Ile Leu Leu Ile
35
            5 1(
     He Leu He Leu Ash Met Thr Ash Ser Ser Ser Arg Tyr
                2(
40
     (2) INFORMATION FOR SEQ ID NO: 221:
            (i) SEQUENCE CHARACTERISTICS.
45
                   (A) LENGTH: 17 amino acids
                   (B) TYPE: amino acid
                   (D) TCPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 221:
50
     Met Ash Glu Leu Leu Phe Phe Phe Phe Phe Phe Phe Leu His Phe
              <u>t</u>
                             1(
     Val
55
     (2) INFORMATION FOR SEQ ID NO: 222:
60
        (i) SEQUENCE CHARACTERISTICS
```

			(xi)	(:	E) T	YPE: CPOL	H: 1 ami: OGY: SCRII	no ac line	cid ear			: 22:	I :			
5	Met :	Lys	Ph€	Thr	Thr !	Leu	Leu	Phe	Leu	Ala 10	Ala	Val	Ala	Gly	Ala 15	Leu
10	Val	Тут	Ala	Glu 2(Asp	Alā	Ser	Ser	Asp 25	Ser	Thr	Gly	Ala	Asp 30	Pro	Ala
	Gln	Glu	Ala 35	Gly	Thr	Ser	Lys	Prc 40	Asn	Glu	Glu	Ile	Ser 45	Gly	Pro	Ala
15	Glu	Fro 50	Ala	Ser	Prc	Pro	Glu 55	Thr	Thr	Thr	Thr	Ala 60	Gln	Glu	Xaa	Sei
20	Ala 65	Ala	Ala	Val	Gln	Gly 70	Thr	Ala	Lys	Val	Thr 75	Ser	Ser	Arg	Gln	Glu 80
	Leu	Asn	Frc	Leu	Lys 8t	Ser	lle	Val	Glu	Lys 9(Ser	Ile	Leu	Leu	Thr 95	Glu
25	Gln	Ala	Leu	Ala 100	Lys	Ala	Gly	Lys	Gly 105	Met	His	Gly	Gly	Val 110	Pro	GJ?.
	Gly	Lys	Gln 115	Ph∈	Il€	Glu	Asn	Gly 120	Ser	Glu	Phe	Ala	Gln 125	Lys	Leu	Leu
30	Lys	Lys 130	Ph∈	Ser	Leu	Leu	Lys 135	Pro	Trp	Ala						
35	(2)	INF		TION SEQU						:						
40				((A) I (B) T (D) T	LENGT TYPE : TOPOI	TH: 5 : ami LOGY:	00 am ino a lin	nino ncid near	ació): 22	3:			
45	Met 1		Gly	Cys	Gly	· Ile	Pro	Ala	Leu	Gly 10		Leu	Leu	Leu	Leu 15	Glr
4 5	Xaa	Ser	Ala	Asp 2(Gly	Asr.	Gly	·Ile	Gln 25	Gly	Phe	Phe	Tyr	Pro 30		Sei
50	Cys	Glu	Gly 35	Asp	∶Il∈	Trp) Asp	Arg 40		Ser	Cys	Gly	Gly 45		Ala	Alé
	Ile	Arg 50														
55	(2)	INF	ORMA	MO ETZ	ı FOF	R SEÇ) ID	N ⊙:	224:							
60				SEQU	JENCI	E CH		reris	TICS	:	dr.					

```
(E) TYPE: amino acio
                 (D) TOPOLOGY: linear
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 224:
     Met Glu Ala Val Phe Thr Val Fhe Phe Phe Leu Leu Phe Cys Phe
      1 5
10
     (2) INFORMATION FOR SEQ ID NO: 225:
           (i) SEQUENCE CHARACTERISTICS.
                  (A) LENGTH: 155 amino acids
                  (E) TYPE: amino acid
15
                  (D) TOPOLOGY: linear
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 225:
     Met Gly Fhe Gly Ala Thr Leu Ala Val Gly Leu Thr Ile Phe Val Let
                                    1(
20
     Ser Val Val Thr Ile Ile Ile Cys Phe Thr Cys Ser Cys Cys Let
     Tyr Lys Thr Cys Arg Arg Frc Arg Frc Val Val Thr Thr Thr Se:
25
                      4(
     Thr Thr Val Val His Ala Pro Tyr Pro Gln Pro Pro Ser Val Pro Pro
30
     Ser Tyr Prc Gly Pro Ser Tyr Gln Gly Tyr His Thr Met Pro Pro Glr.
                             70
     Pro Gly Met Pro Ala Ala Pro Tyr Pro Met Gln Tyr Pro Pro Pro Typ
35
     Pro Ala Gin Pro Met Gly Pro Pro Ala Tyr His Glu Thr Leu Ala Gly
                      10:
     Gly Ala Ala Ala Pro Tyr Pro Ala Ser Gln Pro Pro Tyr Asn Pro Xac
40
                          12(
     Tyr Met Asp Ala Pro Lys Xaa Xaa Ser Glu His Ser Leu Ala Ser Leu
            131
45
     Ala Ala Thr Trp Leu Cys Cys Val Cys Ala Xaa
     145
                     150
50
     (2) INFORMATION FOR SEQ ID NO: 226:
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 10 amino acids
                  (B) TYPE: amino acid
55
                  (D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 226:
     Met Gly Phe Gly Ala Thr Leu Ala Val Gly
       5 10
60
```

ARECORRO NA CERCULARA

فالمناف المناف المناف المناف

```
(2) INFORMATION FOR SEQ ID NO: 227:
 5
            (i) SEDUENCE CHARACTERISTICS:
                   (A) LENGTH: 20 amino acids
                   (B) TYPE: amine acid
                   (D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 227:
10
     Met Ser Ile Phe Leu Val Met Ser Ile Ser Cys Ser Ser Thr Ser His
             Ē
                                      10
     Cys Tyr Ser Ph€
15
     (2) INFORMATION FOR SEQ ID NO: 228:
20
            (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 94 amine acids
                   (E) TYPE: amino acid
                   (D) TCPOLOGY: linear
25
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 228:
     Met Ser Phe Ser Phe Ile Ile Phe Leu Leu Leu Val Cys Gln Glu Ile
                                 10
30
     Thr Phe Cys Met Ser Tyr Gly Asp Ala Val Ash Cys Phe Ser Glu Cy:
     Phe Ser Asn Leu Gln Thr Ile Tyr Ile Ser Cys Leu Gln His Ala Val
35
     Cys Lys His Ser Val Ile Trp Ser Ile Gln Leu Phe Val Arg Ala Lev
     Pro Ile Ser Lys Cys Ala Glu Leu Ser Ile Asp Gly Ile Phe Arg Sei
40
     Phe His Glu Asn Trp Lys Cys Ser Trp Val Ala Pro Thr Xaa
                     85
                            9(
45
     (2) INFORMATION FOR SEQ ID NO: 229:
            (i) SEQUENCE CHARACTERISTICS.
50
                   (A) LENGTH: 94 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 229:
55
     Met Ser Phe Ser Fhe Ile Ile Phe Leu Leu Leu Val Cys Gln Glu Ile
                     Ē
                                10
     Thr Phe Cys Met Ser Tyr Gly Asp Ala Val Asn Cys Phe Ser Glu Cy:
               26 25 36
60
```

	Fn€	Ser	Asn 3 <u>E</u>	L€u	Glr.	Thr	∏€	Tyr 4(11e	Ser	CAE	Leu	Gln 4:	His	Ala	Val
5	Сує	Ly's 50	His	Ser	Vāl	lle	îŗ	Ser	Il€	Glr.	Leu	Ph∈ €(Val	Arq	Ala	Leu
	Prc €:	Il€	Ser	Lys	Cys	Ala 7(Glu	Leu	Ser	lle	Asp 7 <u>t</u>	Gly	Il€	Phe	Arg	Se:
10	Fi₁€	His	Glu	Asn	Trp 85	Lys	Cys	Ser	Trp	Val 90	Ala	Prc	Thr	Χāē		
15	(2)	INF	ORMA'	noi	FOR	SEÇ	ID 1	NC: 2	230:							
20				(A) L B) T D) T	ENGT: YPE: OPOL	H: 3 ama OGY:	7 am nc a lin	inc cid ear	acid		: 23	С.			
25	Met 1	Gly	Trp	Ser	Ala 5	Gly	Leu	Leu	Ph€	Leu 1(Leu	Il€	Leu	Tyr	Leu 15	Pro
4 -	Val	Pro	Gly	Trp 20	Met.	Glu	Arg	Glu	Asp 25	Gly	Gly	Asp	Gly	Thr 3(Ser	Ph€
30	Thr	Ser	Gly 35	Ser	TrŢ											
35	(2)	INF			ENCE	CHA LENGT	RACI H: 8	ERIS	TICS ino		i.s					
40			(xi)		D) I	YPE: OPOL E DE	.CGY :	lir	ısəı	EQ I	D NC	: 23	1:			
	M∈t 1	Alā	. Thr	Leu	Trp	Gly	Gly	Leu	Leu	Arg 10		Gly	Ser	Leu	Leu 15	Sei
45	Leu	Ser	. Cys	Leu 20	Ala	Leu	Ser	Val	Leu 25	Leu	Leu	Ala	His	Val 30	Glr	Thr
50	Frc	Fro	Arg 35	r Ile	Ser	Æ¢	Met	Ser 4(Asp	Val	Asn	Val	Ser 41	Ala	Leu	Pro
50	∏€	Lys 50		∶Il∈	Leu	Gly	Il∈ 5£	- Phe	Ile	Ilε	Arg	Thr 60		leu	Arg	Lys
55	Il∈ €!	Val	. Ile	- Ala	. Phe	Met 70		Trp	Ser	Pro	Cys 75		Cys	Gly	Gly	Leu 80
	Met															

A Francisco Company

	(2)	INFO	ORMAT	ion	FOR	SEÇ	ID N	: 0:2	32:							
5			(i) S	()	A) LI B) T D) T	ENGT: YPE : CPOL	H: 3 ami: OGY:	01 am no ao line	minc cid ear	aci		: 23:				
10	M∈t :	Asp	Ala	Νā	Trr	Trp	Alā	Vāl	Val	Val 10	Leu	Ala	Ala	Phe	Pro 15	Sei
15	Leu	Gly	Ala	Gly 2(Gly	Glu	Thr	Pro	Glu 25	Ala	Pro	Pro	Glu	Ser 30	Trp	The
1.	Gln	L∈u	Trp 35	Ph€	Fhe	Æg	Fhe	Val 40	Val	Asn	Ala	Ala	Gly 4:	Tyr	Ala	Xaa
20	Fh€	Met 50	Val	Fro	Gly	Tyr	Leu 55	Leu	Val	Gln	Tyr	Ph∈ 6(Arg	Arg	Lys	Ası
	Τ <u>Υ</u> Υ	Leu	Glu	Thr	Gly	Arg 7(Gly	Leu	Сув	Ph€	Pro 75	Leu	Val	Lys	Ala	Cys 80
25	Val	Phe	Gly	Asn	Glu E!	Prc	Lys	<i>I</i> .la	Ser	Asp 90	Glu	Val	Pro	Leu	Ala 95	Pro
30	Arg	Thr	Glu	Ala 100	Ala	Glu	Thr	Thr	Pro 105	Met	Trp	Gln	Ala	Leu 110	Lys	Le
	Leu	Phe	Cys 115	Ala	Thr	Gly	Leu	Gln 120	Vāl	Ser	Tyr	Leu	Thr 125	Trp	Gly	Va.
35		130					135					14(
	145		Glu			150					151					16
40			Ala		165					170					175	
45	Pro	Arg	His	Gly 180	Ala	Pro	Met	Tyr	Arg 185	Tyr	Ser	Phe	Ala	Ser 190	Leu	Se
	Asn	Val	Leu 195	Ser	Ser	Trp	Cys	Gln 200		Glu	Ala	Leu	Lys 205	Phe	Val	Se
50	Phe	Pro 210	Thr	Gln	Val	leu	Ala 215		Ala	Ser	Lys	Val 220	Ile	Pro	Val	Me
	Leu 225		Gly	Lys	Leu	Val 23(Arg	Arg	Xaa	Asn 23t	. Glu	His	Trp	Glu	Ту 24
55	Leu	. Thr	- Ala	Thr	Leu 245	∏€	: Ser	Ile	Gly	Val 250		Met	. Phe	Leu	Leu 255	Se
6 0	Ser	Gly	/ Pro	Glu 260	Pro	Ar ç	ser	Ser	Pro 265		Thr	Thr	Leu	Ser 27(Gly	Le

	Il€	Leu	1eu 27!	Ala	Gly	Tyr	$Ii\epsilon$	Ala 280	Fh∈	Asp	Ser	Pr.e	Thr 281	Ser	Asn.	Ίŋ
5	Gìr.	Asp 29(Ala	Cys	Leu	Fro	Il∈ 295	Arg	Cys	His	Æg	300 Cire	Arç			
10	(2)	INF	CRMA'		FOR ENCE	_				:						
1.6				(A) I B) T D) T	YPE: UCGO	ami OGY:	nc a lin	cid ear							
15					TENC											
	Met	S€r	Asp	Leu	Leu	Leu	Leu	Gly	L€u	Ile 1(Gly	Gly	Leu	Thr	Leu 15	L€u
20	L€u	Ьe·u	Leu	Thr 20	Leu	Leu	Ala	Phe	Ala 25	Gly	Tyr	Ser	Gly	Leu 30	Leu	Ala
25	Gly	Val	Glu 3!	Val	Ser	Ala	Gly	Ser 4(Frc	Frc	He	Arg	Asn 45	Val	Thr	Val
25	Ala	Tyr 5(Lys	Ph€	His	Met	Gly 55	Leu	Tyr	Gly	Glu	Thr	Gly	Arg	Leu	Ph€
30	Thr 65	Glu	Ser	Cλ.ε	Ser	11€ 70	Ser	Frc	Lys	L€u	Arg 7:	Ser	Ile	Ala	Val	Ty: 80
	Tyr	Asp	Asn	Pro	His 8t	Met	Val	Pro	Pro	Asp 9(Lys	Cys	Arg	Сле	Ala 95	Vāĵ
35	Gly	S∈r	Il€	Leu 10(Ser	Glu	Gly	Glu	Glu 105	Ser	Pro	Ser	Pro	Glu 110	Leu	Il€
40	Asp	L∈u	Tyr 11t	Glr.	Lys	Ph∈	Gly	Ph∈ 12(Lys	Val	Ph€	Ser	Phe 125	Pro	Ala	Pre
40	Ser	His	Val	Val	Thr	Ala	Thr 135	Phe	Pro	Tyr	Thr	Thr 14(Ile	Leu	Ser	Il€
45	Trp 145	Leu	Alā	Thr	Arg	Arg 150	Val	His	Pro	Ala	Leu 15:	Asp	Thr	Tyr	Ile	Ly: 160
	Glu	Arg	Lys	Leu	Cys 16:	Ala	Tyr	Pro	Arg	Leu 17(Glu	Ile	Tyr	Gln	Glu 17t	Ast
50	Gln	∏€	His	Ph.∈ 180	Met	Che	Pro	Leu	Ala 185	Xaa	Gln	Gly	Asp	Phe 190	Tyr	Vāl
55	Pro	Glu	Met 19:	Lys	Glu	Thr	Glu	Trp 200	Lys	Trp	Arg	Gly	Leu 205	Val	Glu	Alć
	Ile	Asp 21(Thr	Gln	Val	Asp	Gly 215	Thr	Gly	Ala	Asp	Thr 220	Met	Ser	Asp	The
60	Ser 225	Ser	Val	Ser	Leu	Glu 230	Val	Ser	Pro	Gly	Ser 235	Arg	Glu	Thr	Ser	Ala 240

	Ala	The	Leu	Ser	Pro 241	Gly	Ala	Ser	Ser	Arg 250	Gly	Trp	Asp	Asp	Gly. 255	Ast
5	Thr	Arg	:er	Glu 260	His	Ser	Тут	Ser	Glu 26:	Ser	Gly	Ala	Ser	Gly 270	Ser	Se:
	Phe	Glu	Glu 275	Leu	Asp	L∈u	Glu	Gly 280	Glu	Gly	Frc	Leu	Gly 285	Glu	Ser	Arç
10	L∈u	Asp 290	Pro	Gly	Thr	Xaa	Pro 295	Leu	Gly	Thr	Thr	1.ys 30(Trp	Leu	Trp	Glu
15	Pro 305	Thr	Ala	Pro	Glu	Lys 310	Gly	Lys	Glu							
20	(2)	INF	(i)		ENCE (A) I (B) I	CHA LENGI TYPE:	RACI TH: 4 : am:	TERIS 18 am ino a : lir	TICS mino acid mear	acio						
25				SEÇ										Lev	ī en	Ph€
	Pro 1	Gln	Ser	. Leu	11€	· Len	n His	: шеи	. Leu	1(Pile	rne	FILE	Deu	25	1110
30	Let	ı Phe	Ph€	: 11∈ 20		∃le	Phe	e Leu	Phe 25		e Leu	: Gln	Cys	Leu 30	Thr	Ph€
35	Let	ı Phe	35 35		Pro	Arç	i CJ?	Arg 40		His	: Gly	, Let	1 Cys 45	Phe	Lys	: Ph€
40	(2)) INI		ATIO												
			(i)	SEÇ				TERI:			άι					
45								ino			u.					
			(xi) SE				: li IPTI		SEQ	ID N	0: 2	35:			
50	Pr	o Ali	a Le	u Ar	g Pro	o Al	a Le	u Le	u Trj	p Ala		u Lei	u Ala	a Lei	1 Trj 1!	p Lei
	Су	s Cà	s Al	a Th		c Ar	g Me	t Hi	s Cy: 2:		r Va	l Gl	u Me	t Ala		t Asr
55	Pr	c Va	-													
60	(2	.) IN	FORM	OITA	n fo	R SE	Q ID	NO:	23€	:						

5			(3) (x3)	(A) L B) T D) T	ENGT YPE : CPCL	H: 3 ami OGY:	13 a nc a lin	minc cic ∈a:			: 23	€:			
10	Met I	Thr	Arg	Gly	ř. GľÀ	Pro	Gly	G33.	Arç	Prc 1	Gly	Leu	Prc	Gln	Pro Ii	Pre
	Frc	Leu	leu	Leu 20	Leu	Leu	Leu	Leu	Хаа 21	L€u	Leu	Leu	Val	Thr 3(Ala	Gli
15	Pro	Pro	Lys 35	Pro	Ala	Gly	Val	Tyr 4(I3x	Ala	Thr	Ala	Tyr 45	Trp	Met	Pr
	Ala	Glu 5(Lys	Thr	Val	Gln	Val 55	Lγε	As r.	Val	Met	Asp 60	Lys	Asr.	Gly	Asj
20	Ala Et	Tyr	Gly	Phe	Tyr	Aen 70	Asn	Ser	Val	Lys	Thr	Thr	Gly	Trp	Gly	11¢
25	Leu	Glu	∷i∈	Arg	Ala 85	Gly	Tyr	Gly	Ser	Gln 9(Thr	Leu	Ser	Asn	Glu și	Il.
Au al	lì€	Met	Ph€	Val 10(Ala	Gly	Phe	Leu	Glu 1(!	Gly	Тут	Leu	Thr	Ala 11(Prc	Ha:
30	Met	Asr.	Asp 111	His	Tyr	Thr	Asn	Leu 12(172	Fro	Gln	Leu	Il∈ 125	Thr	Lys	Pr
	Ser	11¢ 13/	Met	Asp	Lys	Val	Gln 135	Asp	Phe	Me∙t	Glu	Lys 14(Gln	Asp	Lys	Trj
35	Thr 145	Arg	Lys	Asn	Il€	Lys 150	Glu	Туг	Lys	Thr	Asp 155	Ser	Ph€	Trp	Arç	His 160
40	Thr	Gly	Tyr	Val	Met 165	Ala	Gln	lle	Asp	Gly 17(Leu	Tyr	Val	Gly	Ala 175	Ly:
40	Lys	Yrć	Ala	11e 180	Leu	Glu	Gly	Thr	Lys 185	Pro	Met	Thr	Leu	Phe 19(Gln	11.
45	Gln	Ph€	Leu 199	Asr.	Ser	Val	Gly	Asp 200	Leu	Leu	Asp	Leu	Ile 205	Pro	Ser	Lei
	Ser	Prc 21(Thr	Lys	Asn	Gly	Ser 21!	Leu	Lyε	Val	Phe	Lys 22(Arg	Trp	Asp	Me:
50	Gly 22t	His	СА.	Ser	Ala	Leu 230	Il∈	Lys	Vāl	Leu	Pro 235	Gly	Phe	Glu	Asn	I1. 24(
55	Leu	Ph∈	Ala	His	Ser 245	Ser	Trp	Tyr	Thr	Tyr 25(Ala	Ala	Met	Leu	Arg 25:	Tie
55	Тут	Lys	His	Trp 260	Asp	Phe	Asr.	Xaa	ile 265	Asp	Lys	Asp	Thr	Ser 27(Ser	Sex
60	Arg	Leu	Ser 271	Phe	Ser	Ser	îyr	Pro 280	Gly	Phe	Leu	Glu	Ser 285	Leu	Asp	Asj

		Tyr 29(Il€	Leu :	S∈r S		31y . 295	Leu :	ll∈ :	Leu :	Leu	Gln ' 300	Thr	Thr I	Asn :	Se:
5	Val 305	P'n€	Arn	Lys '		L∈u 1 310	Leu	Lys (Glr.							
10	(2)			NOI												
15				(F	A) LE 3) TY O) TO	NGTH PE: POLC	H: 29 amin XGY:	96 am no ac line	mino rid ear	acio		235	ī:			
20	M∈t 2	Leu	Gln	Gly	Fro t	Gly	Ser	L€u	Leu	Leu 1(Leu	Phe	Leu	Ala	Ser 15	His
20	Сле	Cys	Leu	Gly 2(Ser	Ala	Yrā	Gly	Leu 25	Ph∈	Leu	Phe	Gly	Gln 30	Frc	As]
25	Ph€	Ser	Tyr 35	Lys	Arg	Xaa	Asn	Cys 4(Lys	Pro	Ile	Pro	Val 45	Asn	Leu	Glı.
	Leu	Cys 5(His	Gly	lle	Glu	Tyr 5!	Gln	Asn	Met	Arg	Leu 6(Pro	Asn	Leu	Lev
30	Gly €:	His	Glu	Thr	Met	Lys 70	Glu	Val	Leu	Glu	Gln 75	Ala	Gly	Ala	Trp	11+ 80
25	Frc	Leu	: Val	Met	Lys 85	Gln	Cys	His	Fro	Asp 90	Thr	Lys	Lys	Phe	Leu Ç	C7::
35	Ser	Leu	Ph€	Ala 100	Pro	Val	Cys	L∈u	Asp 105	Asp	Leu	Asp	Glu	Thr 11(lle	Gl:.
40	Fro	Cys	His 115	Ser	Leu	Cys	Val	Gln 12(Val	Lys	Asp	Arg	Cys 125	Ala	Pro	Vā.
	Меt	Ser 13(: Phe	Gly	Phe	Prc 135		Prc	Asp	Met	Leu 140		Cys	Asp	Ar (
45	Ph∈ 145		o Glr	ı Asp	Asn	Asp 150	Leu	Cys	Ile	Pro	Leu 155	Ala	Ser	Ser	Asp	Hi: 160
50	Lev	: Lei	ı Pro	> Ala	Thr 165	Glu	Glu	Ala	Pro	Lys 17(: Val	Cys	Glu	Ala	Cys 17!	Ly:
30	Asr	ı Ly:	s Ası	Asp 180		Asp	Asr	. Asp	Ile 185		: Glu	Thr	Leu	19(Lys	Asr.
55	Asp	o Pho	e Ala 19	a Lev	Lys	Ile	Lys	Val 200		Glu	ı Ile	. Thr	Tyr 205	: Ile	e Asn	Arç
	Ası	p Th 21		s Il€	: Ile	· Leu	Glu 215		Lys	Ser	. Lys	220	īle	е Туг	Lys	: Leu
60	Ası	n Gl	y Va	l Ser	Glu	a Arg	. Asj	: Leu	ı Lys	Lys	s Sei	. Va.	l Let	ı Trp	Leu	ı Ly:

```
23(
                                         235
      Asy Ser Leu Glm Cys Thr Cys Glu Glu Met Ash Asp Ile Ash Ala Fro
                     245 250 251
  5
      Tyr Leu Val Met Gly Gln Lys Gln Gly Gly Glu Leu Val Ile Thr Ser
                                  26!
      Val Lys Arg Trp Gln Lys Gly Gln Arg Glu Phe Lys Arg Ile Ser Arg
 10
           275
                      281
      Ser Ile Arg Lys Leu Gln Cys Xaa
15
      (2) INFORMATION FOR SEQ ID NO: 238:
            (i) SEQUENCE CHARACTERISTICS:
20
                  (A) LENGTH: 92 amino acids
                   (E) TYPE: amuno acid
                  (D) TOPCLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 238:
25
      Met Ala Ser Leu Gly His Ile Leu Val Phe Cys Val Gly Leu Leu Thr
      Met Ala Lys Ala Glu Ser Fro Lys Glu His Asp Pro Phe Thr Tyr Asp
                20 25
30
      Tyr Gln Ser Leu Gln Ile Gly Gly Leu Val Ile Ala Gly Ile Leu Ph\epsilon
      lie Leu Gly Ile Leu Ile Val Leu Ser Arg Arg Cys Arg Cys Lys Phe
35
      Ash Gln Gln Gln Arg Thr Gly Glu Pro Asp Glu Glu Glu Gly Thr Phe
                        7(
40
     Arg Ser Ser Ile Arg Arg Leu Ser Xaa Arg Xaa Arc
                    ۶L
45
     (2) INFORMATION FOR SEQ ID NO: 239:
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 71 amine acids
                   (B) TYPE: amino ació
50
                  (D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 239:
     Met Pro Gly Thr Phe Leu Arg Pro Phe Val Phe Leu Phe Leu Phe II\epsilon
                            2.0
55
     Cys Cys Cys Leu His Ser Gly Gly Leu Gly Gly Val Pro Leu Pro Pro
     Phe Pro Pro Gln Ala Gln Arg Gly Glu Gly Pro Gly Lys Trp Met Sei
60
        35 4(
```

THE STATE OF THE S

```
Fro Fro Leu Pro Pro His Pro Val Val Ala Pro Pro Thr Pro Ser Pro
                          6.5
     Ser Arg Gly Cys Val Leu Leu
     (2) INFORMATION FOR SEC ID NO: 240:
10
             (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 71 amino acids
                    (E) TYPE: amino acid
                    (D) TOPOLOGY: linear
15
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 240:
      Met Pro Gly Thr Phe Leu Arg Pro Phe Val Phe Leu Fhe Leu Phe Ile
20
      Cys Cys Cys Leu His Ser Gly Gly Leu Gly Gly Val Pro Leu Pro Pro
      Phe Pro Fro Gln Ala Gln Arg Gly Glu Gly Pro Gly Lys Trp Met Ser
25
      Pro Pro Leu Pro Fro His Pro Val Val Ala Pro Pro Thr Pro Ser Pro
                                          €(
                              55
30
      Ser Arg Gly Cys Val Leu Leu
      (2) INFORMATION FOR SEQ ID NO: 241:
35
             (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 28 amino acids
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
40
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 241:
      Met Phe Tyr Val Leu Ser Val Ser Xaa Leu Xaa Leu Phe Leu Ala Cys
                                          10
45
      Gly Leu Cys Leu Xaa Leu Leu Thr Gly Lys Leu Le:
                                      25
                   20
 50
       (2) INFORMATION FOR SEQ ID NO: 242:
              (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 58 amino acids
 55
                     (E) TYPE: amino acid
                     (D) TOPOLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 241:
       Met Lys Leu Phe Asp Ala Ser Pro Thr Phe Phe Ala Phe Leu Leu Gly
 60
                                           10
```

34-

```
His Ile Leu Ala Met Glu Val Leu Ala Trp Leu Leu Ile Tyr Leu Leu
                          25
      Gly Fro Gly Trp Val Pro Ser Ala Leu Xaa Arg Leu His Fro Gly His
      Leu Ser Gly Ser Val Leu Val Ser Ala Ala
10
      (2) INFORMATION FOR SEQ ID NO: 243:
15
           (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 123 amino acios
                   (E) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEC ID NO: 243.
20
     Met Ile Leu Gly Gly Ile Val Val Val Leu Val Phe Thr Gly Phe Val
                                       1(
     Trp Ala Ala His Asr. Lys Asp Val Leu Arg Arg Met Lys Lys Arg Ty:
25
     Fro Thr Thr Fhe Val Met Val Val Met Leu Ala Ser Tyr Phe Leu Il\epsilon
                       40
30
      Ser Met Phe Gly Gly Val Met Val Phe Val Phe Gly Ile Thr Phe Pro
                            25
     Leu Leu Met Phe Ile His Ala Ser Leu Arg Leu Arg Asn Leu Lys
35
     Ash Lys Leu Glu Ash Lys Met Glu Gly Ile Gly Leu Lys Arg Thr Pro
     Met Gly Ile Val Leu Asp Ala Leu Glu Glu Glu Glu Gly Ile Asr.
40
                       105
     Arg Leu Thr Asp Tyr Ile Ser Lys Val Lys Glu
        11: 120
45
     (2) INFORMATION FOR SEQ ID NO: 244:
            (i) SEQUENCE CHARACTERISTICS:
50
                  (A) LENGTH: 73 amino acid:
                   (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 244:
55
     Ala Leu Val Ser Gly Gln Leu Cys Met Glu Ile Ala Arg Gly Asn Il\epsilon
      1 5 1(
     Phe Phe Leu Ash Xaa Leu Val Thr Thr Phe Cys Cys Ser Cys Leu Leu
60
```

	Leu	Ser	Vā.] 3:	Хаа	יבילI	Leu	Eis	Xaa 40	GŢŻ	Ph∈	Ph∈	J.Ž.L.	S∈r 45	Ser	Leu	СŻъ
5	Lys	Cys 5(C7.E	Ph€	Val	L€u	Val 55	Val	Leu	Ser	Arg	Ile 6i	Gly	Ser	Val	Ast.
	Glu 65	Thr	Trp	Ser	CÀE	Asn 7(Phe	Ser	Il€							
10																
	(2)	INF	ORMAT	NOL	FOR	SEÇ	ID N	10: 2	245:							
15				(1	A) LI B) T D) T	ENGT YPE: OPOL	H: 4 amin OGY:	9 am no a lin	ino d ció ear	acid		: 245	á:			
20	Thr 1	Pro	Ala	Thr	mhr.	Ser	Ser	Ser	Ser	Ser 10	Pro	Leu	Ph∈	Leu	Ser 15	Ser
25	Frc	Asp	Trp	Ser 20	Ser	Cys	Prc	Ser	Gly 25	Ser	СЛЕ	Ile	Ala	Pro 30	Trp	Суя
23	Thr	Hiε	Trp 3:	Ser	Ser	Ile	Leu	Pro 40	Ser	Leu	Xaa	lle	Thr 45	Ser	Ser	Il€
30	Prc															
35	(2)	INF		TION SEÇU	ENCE	СНА	RACT	ERIS	TICS							
							TH: 3 : ami			aci	Ġ!					
40			(xi)) SEÇ			LOGY: ESCRI			EQ I	D NC	: 24	6 :			
	Met 1	Ala	a Arg	Vāl	Fro	Prc	Leu	Ser	Ser	Ser 10	Trp	Thr	Ser	Ser	Arg 15	Ту
45	Arg	Arg	g Trp	Leu 2(Cys	Cys	Pro	Val	Trp 25	Trp	Thr	Thr	Phe	Trp 30		Th
50	Ala	Trp	Ser 3:	Leu	Thr	Lys	His	Leu 40		Lys	Asp	Vāl	Thr 45	Asp	Ala	1 1
50	Arg	Asp 50		His	Val	Lys	61y 55		Met	Tyr	Gln	Trp	Ile	Glu	. Gln	As
55	Met 6:	Gli	u Lys	: Tyr	Ile	Leu 7(Gly	' Asp	Glu	Thr 7:	Phe	Ala	Val	Leu	S∈ 8
	Arg	Lei	u Val	Ala	His 85	Gly	y Lys	Gln	Leu	Phe 9(Leu	lle	Thr	Asn	Ser 95	Pr
60	Phe	، دد	r Phe	- Val	Asn	LVS	: Glv	Met	Arc	His	Met	Val	Glv	Pro	Asp	Tr

				300					105					110		
5	Arç	His	Ser II:	Ser	Met	Trp	Ser	Leu 120	Ser	Arç	Gin	Thr	Ser 125	Prc	Ala	S€1
3	Ser	leu 13(Thr	Gly	A.ē	Thr	Ph∈ 135	Arç	Lys	Leu	Asp	Glu 14(Lys	Gly	Ser	Le:
10	Glm 145	Trp	Asp	Æģ	Il€	Thr 15(Arg	Leu	Glu	Lys	Gly 15t	Lys	Il€	Tyr	Arç	Gl:. 160
	Gly	Aer.	Leu	Ph€	Asp 165	F'n€	L€u	Arç	Leu	Thr 17(Glu	Trp	Arg	Gly	Frc 17:	Arç
15	Val	L€u	Tyr	Fh∈ 180	Gly	Asp	His	Leu	Tyr 18:	Ser	Asp	Leu	Ala	Asp 190	Leu	M€.
20	Leu	Arg	His 195	GΣγ	Trp	Arç	Thr	Gly 200	Ala	Il€	īiε	Pro	Glu 201	L∈u	Glu	Arı
20	Glu	11∈ 21(Αrq	Il€	$\mathbb{I}1\epsilon$	Asn	Thr 215	G] u	Gln	Tyr	Met	His 220	Ser	Leu	Thr	Ίŋ
25	Gln 225	Gin	Ala	Leu	Thr	Gly 23(Leu	Leu	Glu	Αrç	Met 235	Gln	Thr	Tyr	Gln	As ₁ 240
	Ala	Glu	Ser	Arç	Gln 245	Val	L∈u	Ala	Alā	Trp 25(Met	Lys	Glu	Arg	Gln 255	GÎ.
30	Leu	Arç	Chē	∏€ 26(Thr	Lys	Ala	Leu	Phe 26:	Asn	Ala	Gln	Phe	Gly 27(Ser	13+
35	Ph€	Arç	Thr 27!	Ph∈	His	Asn	Prc	Thr 28(Tyr	Phe	Ser	Arg	Arg 285	Leu	Val	Ar:
J.	Phe	Ser 29(Asj.	Leu	Tyr	Met	Ala 295	Ser	L∈u	Ser	Сув	Leu 30(Leu	Asn	Tyr	Arc
40	Val 305	Asp	F'n€	Thr	Phe	Tyr 310	Prc	Arg	Arg	Thr	Pro 315	Leu	Gln	His	Glu	Al: 321
	Pro	Leu	1rp	M⊕t	Asp 325	Gln	Leu	Leu	His	Arg 33(Leu	His	Glu	Asp	Pro 335	Leu
45	Pro	Trp	Хаε													
50	101	71.00	OD147	T: T S . T . T	FOE	cro	770	vic. '	7 4 5 .							
50	(4)	3147	ORMA: (i)	SEÇU	ENCE	CHA	RACT	ERI <i>S</i>	TICS							
55			(xi)	((A) I (B) T (D) T UENC	YPE:	ami :OGY	nc a lin	cic ea:			: 24	7 :			
60	Met :	Alā	Leu	Leu	Ser 5	Суғ	Val	Vāl	Asp	l(Ph∈	Leu	Gly	His	Ser 11	Leu

A STATE OF THE STA

Xaa Va.

5																
	(2)	INFO	CRMAT	NOIT	FOR	SEÇ	ID 1	NC: 2	248							
10			(i`) (xi)	(A) L E) T D) T	ENGT YPE: OPGL	H: 3 ami OGY:	39 a nc a lin	manc cic ear	aci		: 24	٤:			
15	M∈ t	Asn	Trp	Glu	Leu £	Leu	Leu	Trp	L€u	Leu 10	Val	Leu	Cys	Ala	Leu 15	L€ι
	L€u	Leu	Leu	Val 20	Gln	L€u	Leu	Arg	Ph∈ 25	Leu	Arg	Ala	Asp	Gly 3(Asp	Lev
20	The	Leu	Leu 3i	Trp	Ala	Glu	Trp	Gln 4(GJ?.	Arg	Arg	Pro	Glu 41	Trp	Glu	Leu
25	Thr	Asp 5(Met	Val	Vāl	Trp	Vāl 55	Thr	Gly	Alā	Ser	Ser 60	Gly	∏e	Gly	Glu
25	Glu 6t	Leu	Ala	Tyr	Gln	Leu 70	Ser	Lys	Leu	Gly	Val 75	Ser	Leu	Val	Leu	Se:
30	Ala	Arg	Arg	Val	His 85	Glu	Leu	Glu	Arç	Val 9(Lys	Arg	Arg	Cys	Leu 9!	Gli
	Asn	Gly	Asn	L∈u 100	Lys	Glu	Lys	Asp	:1e 10!	Leu	Val	Leu	Pro	Leu 110	Asp	Lev
35	Thr	Asp	Thr 115	Gly	Ser	His	Glu	Ala 12(Ala	Thr	Lys	Ala	Val 125	Leu	Gln	GI:
40	Ph∈	Gly 13(Arç	∏e	Asp	lle	Leu 135	Val	Asn	Asn	Gly	Gly 14(Met	Ser	Gln	Arş
40	Ser 145	Leu	Cys	M∈t	Asp	Thr 150		Leu	Asp	Val	Tyr 155	Arg	Lys	Leu	Ile	Gl: 16(
45	Leu	Asn	Tyr	Leu	Gly 165	Thr	Vāl	S∈r	Leu	Thr 17(Lys	Cys	Val	Leu	Pro 17!	His
	Met	Ile	Glu	Arg 18(Lys	Gln	Gly	Lys	Ile 185	Val	Thr	Val	Asn	Ser 190	Ile	Le.
50	Gly	Ξle	11∈ 195	Ser	Val	Pro	L€u	Ser 20(Il€	Gly	Tyr	Cys	Ala 205	Ser	Lys	His
EE	Ala	Leu 210	λrg	Gly	Phe	Phe	Asn 215	Gly	Leu	Arg	Thr	Glu 220	Leu	Ala	Thr	Туі
55	Pro 22!	Gly	ïle	IJ€	Val	Ser 230	Asn	∏e	Cys	Pro	Gly 23!	Pro	Val	Gln	Ser	As: 24(
60	IJ€	Val	Glu	Asn	Ser 245		Ala	Gly	Glu	Val 25(Thr	Lys	Thr	Ile	Gly 251	Ası

	Ast.	eĵλ.	Asp	Glr. 260	Ser	His	Lys	Met	Thy 265	Thr	Ser	Arg	Cys	Val 27(Arg	Leu
5	Met	Leu	11€ 275	Ser	Met	Ala	Ast.	Asp 28(Leu	Lys	Glu	Val	Trp 285	Il∈	Ser	Glu
10	Glr.	Frc 29(Ph€	Leu	Leu	Val	Tha 29!	Tyr	L∈u	Imp	Gln	Tyr 300	Met	Prc	Thr	TrŢ
10	Ala 301	Trp	Trp	∏€	Thr	Asn 31(Lys	M∈t	Gly	Lys	Lys 315	Arg	He	Glu	Asn	Ph€ 320
15	Lys	Ser	Gly	Val	Asp 325	Ala	Asp	Ser	Ser	Tyr 33(Phe	Lys	Ile	Ph∈	Lys 33t	The
	Lys	His	Ası													
20																
	(2)	I MF (NC: : ERIS		:						
25				(A) L B) T D) T	ENGT YPE : OPCL	H: 9 amı OGY:	e am nc a lin PIIO	ino cić ea:	acid		: 24	9 :			
30	Met	Gly	Ala	Arg	Pro	Gly	Gly	His	Prc	Gln 1(Lys	Trp	Ser	Phe	Leu 1!	Tr
35	Ser	Leu	Ala	Leu 20	Trp	Leu	Pro	Leu	Ala 25	Leu	Ser	Val	Ser	Leu 3(Phe	Lev
J.	Gly	Leu	Ser 3t	Leu	Ser	Pro	Pro	Gln 4(Fro	Gly	Leu	Ser	Leu 45	Irp	Càt	The
40	Leu	£er 5(Ţγr	Сув	Cys	Glu	Gln 55	ЧŢ	Lys	Phe	Lys	Gly 60	Thr	Pro	Ser	Pre
	Ala (!	Leu	Leu	Asn	Leu	Gly 70	Thr	Gln	Pro	Lys	Lys 75	Asp	Lys	Lys	Leu	Glu 8(
45	Asp	Ser	Il∈	Ala	Thr 85	Gln	Leu	Arç	Хаа	Leu 90	Pro	Glu	Lys	Asn	Ser gr	Asn
50																
	(2)	INF	ORMA!	TION	FOR	SEQ	ID	NO:	250.							
55			(i)		(A) I	ENGI YPE :	TH: T	ERIS 79 am ino a : lir	nino cić		is					
60			(xi)					PTIC		EQ I	D NC	: 25	C :			

the state of the same of

	Met A	Ala	Leu	Thr	Ph∈	Leu	L←u	Val	Leu	Leu 1(Thr	Leu	Ala	Thr	Leu 1!	C).;
5	Thr 1	Ārģ	L€u	His 20	Arg	Asn	Fh€	Arq	Arg 2:	Gly	Glu	Ser	lle	Tyr 3(Trp	Gly
	Frc '	Thr	Ala 35	Asp	Ser	Gln	Asp	Trar 4(Val	Ala	Ala	Vāl	Leu 4!	Lys	Αrç	Arc
10	Leu l	Leu 5(Gln	Prc	Ser	Arg	Arg 51	Val	Lys	Arg	Ser	Arg 60	Arq	Arg	Frc	Хає
15	Xaa 65	Pro	Pro	Thr	Prc	Asp 7(Ser	Gly	Pro	Glu	Gly 75	Glu	Ser	Ser	Glu	
20	(2)	INF	(i)	(ENCE A) L B) T D) T	CHAI ENGT YPE:	FACT E: 3 ami OGY:	ERIS 54 æ no æ lin	TICS mino cić near	aci		·: 25	1:			
25	Met 1	Gly		ser Ser										Ser	Trp 15	Sei
30	Gly	Pro	: Le≀	ı Gln 20	Gly	Gln	Gln	His	His 25	Leu	Val	Glu	Tyr	Met 3(Glu	Arç
	Arg	Let	ı Ala 35	a Ala	Leu	Glu	Glu	Arg 4(Ala	Gln	. Cys	Gln 45	Asp	Gln	Sei
35		5(-	s Ala			5.5					60				
40	€£					7(75					3ľA 1
				e Ser	85					90					9:	
45				100)				105					110		Lys
			11	Ē				120	0				125			ı Lys
50		13	(13	c.				140	-			g Ser
55	145					150					15					r Lys 160
					16					17	C				17	
60	Ast	ı As	p Th	ar Al 18		e Va	l Ph	e Pr	o Ar		u Ar	g As	p Ph	e Th 19	r Le (u Ala

	Met	A. ē.	Ala 195	Arc	Lys	Ala	£62	Arg 200	Vā.	Arç	Val	Pro	Ph∈ 2C5	Frc	Trp	Val.
5	GŢλ	Tnr 210	Gly	Gln	Leu	Val	177 215	Gly	Gly	Phe	Leu	Tyr 22′	Ph∈	Ala	Arg	Arg
1(+	Frc 225	Frc	Gľy	Arg	Frc	Gly 23(Gly	Gly	Gly	Glu	Met 235	Glu	Asr.	Thr	Leu	Gìr. 240
	Leu	∏e	The	Ph€	His 245	Leu	Alē	Asr.	Arg	Thr 250	Vāl	Val	Yet	Ser	Ser 255	Val
15	Ph€	Frc	Ala	Glu 260	Gly	L€u	∏€	Frc	Pro 265	Tyr	Gly	Leu	Thr	Ala 27(Asp	Thi
	Tyr	Il€	Asp 275	Leu	Alā	Ala	Asp	Glu 280	Glu	Gly	Leu	Trp	Ala 28:	Val	Tyr	Ala
20	Thr	Arg 29(Glu	Asp	Yet	Arç	His 291	L€u	ርን፡ε	Leu	Ala	Lys 30(Leu	Asp	Pro	Glr.
25	Thr 305	≟e·u	Asp	Thr	Glu	Gln 31:	Gln	Imp	Asp	Thr	Pro 31!	Сує	Frc	Arg	Glu	Asr. 320
	Ala	Glu	Ala	Ala	Ph∈ 32!	Хаа	Il€	Сув	Gly	Thr 330	Leu	Tyr	Val	Val	Tyr 335	Asr.
30	Tim	Arg	Prc	Ala 34(Ser	Αrç	Ala	Arg	11∈ 345	Gin.	CAE	Ser	Ph€	Asp 350	Ala	Sei
	Gly	Pro														
35	Gly	Pro														
35			ORMA'													
3 <i>5</i> 40			(i)	SEQUI (ENCE A) L B) T D) T	CHA ENGT YPE: OPCL	KACT E: 1 ami OGY:	ERIS 09 a no a lin	TICS mino cid ear	aci		: 25				
	(2)	INF	(i)	SEQUI () (SEÇI	ENCE A) L B) T D) T UENC	CHA ENGT YPE: OPCL E DE	RACT H: 1 ami OGY: SCKI	ERIS' 09 a no a lin PTIO	TICS mino ció ear N: S	aci EÇ I	D NO			Val	Pro 15	Ser
40 45	(2) Met	INF((i) (xi)	SEQUI ((SEQ Ile	ENCE A) L E) T D) T UENC Asn	CHA ENGT YPE: OPCL E DE	RACT H: 1 ami OGY: SCRI Thr	ERIS' 09 a no a lin PTIO	TICS mino cic ear N: S Pro	aci EQ I Arg 10	D NO Pro	Leu	Fro		15	
4(Met 1	1NFC	(i) (xi) Cys	SEQUI ((SEQ Ile Cys 20	ENCE A) L B) T D) T UENC Asn [CHAMENGT YPE: OPCL E DE Gly Ile	HACT: E: 1 ami OGY: SCKI Thr	ERIS' 09 a no a lin PTIO Thr	rics mino cid ear N: S Pro	e aci EQ I Arg 10 Phe	D NO Pro Lys	Leu Asn	Pro	Trp	15 Lys	Glr.
40 45	Met 1 Prc Arg	1NFC	(i) (xi) Cys Gly Leu	SEQUI ((SEQ Ile Cys 20	ENCE A) L B) T D) T UENC Asn E Met	CHAMENGT YPE: OPCL E DE Gly Tie	RACTI E: 1 ami OGY: SCK1 Thr	ERIS' C9 a no a lin PTIO Thr Phe Gly 40	TICS mino cid ear N: S Pro Fhe 25	EQ I Arg 10 Phe	D NO Pro Lys Pro	Leu Asn	Pro	Trp 3(15 Lys Leu	Glr.
40 45 50	Met 1 Proc Arg	INFO Leu Fhe Leu 50	(i) (xi) Cys Gly Leu 35	SEQUING ((() () () () () () () () (ENCE A) L E) T D) T D) T E) G E G Ser	CHAMENGT YPE: OPCL Gly Ile Trp Leu	HACTHER 1 ami OGY: SCKI Thr Fhe	ERIS' C9 a no a lin prico Thr Phe Gly 40	TICS mino cic ear N: S Pro	EQ I Arg I(Phe Arg Pro	D NO Pro	Leu Asn Ile Pro 60	Pro Pro His 45 Leu	Trp 3(Leu Pro	lis Lys Leu Cys	Gly Ala

WO 98/39446

			£ ;					9(ð.f	
5	Lys Glu L	€u Gl [.] 10		ੁਕਰ ⊖	sly A		Of As I	rp G	ilu F	rc S	·E:			
10	(2) INFOR		N FOR S DUENCE ((A) LE (E) TY	CHAF NGTH	ACTE: : 45	RISTI ami	ICS: nc a	ciás						
15	(> Met Phe T		(D) TO EQUENCE	POLO DES	GY: CRIP	linea TION:	ar : SE					Prc '	Trp . 15	Ala
20	Fhe Arg I		er Thr	Leu :	Phe S	Thr I	lle: 25	Ile:	Ser '	Irp (Ser (31u . 30	Asp	Ser
9.5	Asn Asn S	er Gl 35	lm Val	Tyr :	Met 2	Asn (40	Cys '	Val (Cys	Ser :	Ph€ 45			
25	(2) INFO	21.2 E.T.	on EoE	CEO	יו חד	O • 21	54.							
30	(i) SE	QUENCE (A) LI (B) T (D) TO	CHAF ENGTI YPE: OPOLO	ACTE H: 31 amir XGY:	RIST 5 am no ac line	ICS: nino cid	acio		: 254	:			
35	Met Ala	Gly G	ly Arg	CAE	Gly	Prc	Xaa	Leu 1(Thr	Ala	Leu	Leu	Ala 15	Ala
40	Trp Ile		la Val 20	Ala	Ala	Thr	Ala 25	Gly	Pro	Glu	Glu	Ala 30	Ala	Let
40	Pro Pro	Glu G 31	ln Ser	Arg	Val	Gln 40	Prc	Met	Thr	Ala	Ser 45	Asn	Trp	The
45	Leu Val 50	Met G	slu Gly	Glu	Trp 55	Met	Leu	Lys	Ph€	Tyr 6(Ala	Pro	Trp	Cys
	Pro Ser 65	Cys (Sln Gln	Thr 70	Asp	Ser	Glu	Trp	Glu 75	Ala	Phe	Ala	Lys	Asn 38
50	Gly Glu	Ile I	Leu Gln 81	le	Ser	Val	Gly	Lys 9(Val	Asp	Val	Ile	Gln 95	Glu
F. F.	Pro Gly		Ser Gly 100	Arg	Phe	Phe	Val 105	Thr	Thr	Leu	Pro	Ala 110	Phe	Ph∈
55	His Ala	Lys 2	Asp Gly	, lle	Phe	Arg 120	Arg	Tyr	Arg	Gly	Pro	Gly	Ile	Ph€
60	Glu Asp 130	Leu (Glr. Asr	. Tyr	11€ 135	Leu	Glu	Lys	Lys	Trp 14(Gln	Ser	Val	. Gl:

	Prc 145	Leu	Til	G ₂ y	TT	Lys 150	Ser	Fre	A.ē	Ser	Leu 15t	Thr	M∈t.	Ser	Gly	M∈1
5	Ala	Gîy	Leu	Frie	Ser 165	Il€	Ser	GI?.	Lys	Il∈ 17(Ϋ́τ	His	Leu	His	Asn 175	Ty:
10	Ph∈	Thr	Vā:	Thr 180	Leu	Gly	Il€	Pro	Ala 185	Trp	Cys	Ser	Tyr	Val 190	P'n€	Pr.
	Val	Il€	Ala 195	Thr	Leu	Val	Phe	Gly 200	Leu	Phe	Met	Gly	Leu 205	Val	Leu	Val
15	Val	11e 21(Ser	Glu	САЕ	Phe	Tyr 215	Vāl	Frc	Leu	Frc	Arg 22(His	Leu	Ser	Glu
	Arç 225	Ser	Glu	Gln	Asn	Arg 230	Arg	Ser	Glu	Glu	Ala 235	His	Arg	Ala	Glu	Gl: 240
20	L€u	Glr	Asp	Ala	Glu 245	Glu	Glu	Lys	Asp	Asp 25(Ser	Asn	Glu	Glu	Glu 25:	As:
25	Lys	Yet	S€I	Leu 26(Va]	Asp	Asp	Glu	Glu 261	Glu	Lys	Glu	Asp	Leu 27(Gly	Asj
	Glu	Yet	Giu 27:	Ala	Glu	Glu	Glu	Glu 28(Glu	Glu	Asp	Asn	Leu 285	Ala	Ala	GIA
30	Vāl	Asp 29(Gìu	Glu	Arg	Ser	Glu 295	Ala	Asn	Yet	Gln	Gly 30(Pro	Pro	Gly	Gl:
	Asp 305	Gly	Val	Thr	Arg	Glu 310	Xaa	Ser	Arg	Ala	Хаа 31!					
35	(2)	INF	ORMA!	NOLI	FOR	SEQ	ID I	NO: 1	25E:							
40			(i) .	(A) L B) T D) T	ENGT YPE: OPOL	H: 5 ami OGY:	3 am no a lin	inc cic ear	acid		: 25	5:			
45	Met 1	Leu	Lys	Ala	Leu ţ	Ph€	Arg	Thr	Leu	Gln 1(Ala	Met	Leu	Leu	Gly 15	Va.
50	Trp	ll∈	Leu	Leu 2(Leu	Leu	Ala	132	Leu 2i	Ala	Pro	Leu	Trp	Leu 30	Tyr	СЪt
50	Trp	Arg	Met 35	Phe	Pro	Thr	Lys	Gly 4(Lys	Arg	Asp	Gln	Lys 4t	Glu	Met	Let
55	Glu	Val 50	Ser	Glу	Iĭ€											
60	(2)	INFO	ORMAI	NOI	FOR	SEÇ	ID 1	NO: 2	56 :							

	(i) SEQUENCE CHARACTERISTICS (A) LENGTH: 93 amino acido (E) TYPE: amino acido	
5	(D) TOPOLOGY: linea: (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 256:	
	Met lie His Leu Gly His lie Leu Phe Leu Leu Leu Leu Pro Val Ala	
10	Ala Ala Glr. Thr Thr Pro Gly Glu Arg Ser Ser Leu Pro Ala Phe Ty: 20 25 30	
1.5	Prc Gly Thr Ser Gly Ser Cys Ser Gly Cys Gly Ser Leu Ser Leu Prc 3! 40 45	
15	Leu Leu Ala Gly Leu Val Ala Ala Asp Ala Val Ala Ser Leu Leu IIe 50 55 60	
20	Val Gly Ala Val Phe Leu Cys Ala Arg Pro Arg Arg Ser Pro Ala Glr.	
	Asp Gly Lys Val Tyr Ile Asm Met Pro Gly Arg Gly Xaa 85 90	
25		
	(2) INFORMATION FOR SEQ ID NO: 257: (i) SEQUENCE CHARACTERISTICS:	
30	(A) LENGTH: 12 aminc acid: (B) TYPE: amino acic (D) TOPOLOGY: linea: (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 257:	
35	Pro Gly His Leu Leu Pro His Lys Trp Glu Asn Cys	
40	(2) INFORMATION FOR SEQ ID NO: 25%.	
45	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1852 base pair: (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 258:	
50	TGGCATCTGI GAGCAGCTGC CAGGCTCCGG CCAGGATCCC TTCCTTCTCC TCATTGGCT	60
	ATGGATCCCA AGGGGCTCCT CTCCTTGACC TTCGTGCTGT TTCTCTCCCT GGCTTTTGG	120
55	GCAAGCTACG GAACAGGTGG GCGCATGATG AACTGCCCAA AGATTCTCCG GCAGTTGGGA	180
55	AGCANAGTGC TGCTGCCCCT GACATATGAA AGGATAAATA AGAGCATGAA CAANAGCATC	240
	CACATTOTOG TOACAATGGO AAAATCACTG GAGAACAGTG TOGAGAACAA AATAGTGTO	30(
60	CTTGATCCAT COGAAGCAGG CCCTCCACGT TATCTAGGAG ATCGCTACAA GTTTTATCTC	360

	SPRINGER AD	mateggar.	ACCENALAGO	AFRAAGGAGG	ATGA GGGATG	GTACCTTIATV-	42(
5	ACCCTGGAGA	AAAATGTTTC	ART TOAGORY	TTTTGUCTGC	AGTTGAGGCT	TTATGAGCAG	480
	GTOT COACTO	CAGAAATTAA.	ASTTTTAAAS	AAGACTCAGG	AGAACGGGAC	OT(#BACOTTY)	540
	ADAPTOGGCI	GUACAGTISSA.	GAJAGGERGAG	CATGIGGGTI	acagettguag	TGAAAA(960)	600
10	COADDACEA	CACTGAACCC	APTCAACA(#T	TOCCACCTCC	TGTCCCTCAC	CCTOR ROCO!	660
	CASUATOTO	ACAATATOTA	CATIOTIGUALIO	CTGAGDAACC	CTATCAGCAA	CAATTOCCAG	720
15	ACCTTCAGA	OGTGGCCCG:	ATYGCAGGACA	GACCCCTCAG	AAACAAAACC	ATGGGCAGTG	780
	TAIGOTGGG	TGTTAGGGG:	THYCATCATC	ADTADIOTEA	TGGTGGTAAT	ACTACAGTT	840
	AGAAGAAGAG	GTAAAACGAA	CUMMIACCAS	ALAACAGTGG	PAAAAAAA	CCTTACGATC	90(
20	CACHOCHICA	TOCAGAAACT	apptgaca n	CAMBATCAGA	CTTCGGACTI	COTAATOTTA	960
	AGGATGADUT	TATTTTGAAA	TIPOTA ATOTI	GACATOTOTG	AAGACCTITA	TTUAAATAAA.	1020
25	GTCACATTTT	GACATTOTGO	GA/9BGGCT) +G	RODERBOORER	GGGTGATGTG	GAG0G0GGG1	1080
-	0603603363	CTGCCTGGCC	GBTGCTBT #4	GRECTICETO	TGGIGCTNITT	AGTGCCGGGC	1140
	@33-33 3 5500@	CCAAGACCGS	TGTGGAGTTO	GINGACTGCGG	GTOGGTGTTG	AAGCTGCTCA	1200
30	ACCACCATA	COGGTGGGGT	TGCACTCGCA	CGACATCAAA	TACGGATTCG	GCAGCGGCCA	1260
		ACCGGCGTAG					1320
35						GAGGTCACAC	1380
		GEGCAAGAAC					1440
		TGCCAAAGGG					1500
40						CCACCATON	1560
						GIGGGCAGCA	162
45		GCATGCCCAG					168
						GTGTGGATGG	174
••						A GAGACTTTC	180
50	GCTTTGTAGG	GGTCCTCAAG	- 1960 Calada (19	IATTAAAGAAT	GTTGGTCTAT	G?	185

55 (2) INFORMATION FOR SEC IE NO: 255 -

(i) SEQUENCE CHAFACTEFISTICS:

(A) LENGTH: 371 amino acids

(B) TYPE: amano acic

60 (D) TOPOLOGY: linear

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 259:															
										Gln				Ala	Ph∈	Let
5	Met 1	(- <u>.</u> L	neu	Giu	ř.	wet	F: 2 C:	G.,	ne p	1(1112				15	
5	Leu	Glu	Glu	Ser 20	Gly	Asp	L€u	Gly	Thr 25	Ala	Pro	qzA	Glu	Ala 30	Val	Arç
10	Ala	Fre	Leu 35	Asp	Trp	Alā	L€u	Frc 4(Leu	ser	Glu	Val	Frc 45	Ser	Asp	ıŋ
	Glu	Val 50	Asp	Asp	Leu	Leu	£; CAE	Ser	Leu	Leu	Ser	Pro 60	Pro	Ala	Ser	Leu
15	Asr. Et	Iì€	Leu	Ser	Ser	Ser 70	Asn	Frc	СУЕ	Leu	Val 75	His	His	Asp	His	rdT 38
20	ůλα	Ser	Leu	Prc	Arg 85	Glu	Thr	Val	Ser	Met 90	Asp	Leu	Glu	Ser	Glu 91	Sei
	CAE	Arg	Lys	Glu 100	Gly	Thr	Gln	M∈t	Thr 105	Fro	Gln	His	Met	Glu 11(Glu	Leu
25	Ala	Glu	Gln 115	Glu	Ile	Ala	Arg	Leu 120	Val	Leu	Thr	Asp	Glu 125	Glu	Lys	Sei
	Leu	Leu 13(Glu	Lys	Glu	Gly	Leu 135	Il€	L€u	Pro	Glu	Thr 140	Leu	Pro	Leu	The
30	Lys 145	Thr	Glu	Glu	Gln	11∈ 150	Leu	Lys	Arg	Val	Arg 155	Arg	Lys	ll€	Arg	Asr. 160
35	Lys	Arg	Ser	Ala	Gln 165	Glu	S€r	Arg	Arg	Lys 170	Lys	Lys	Val	Tyr	Val 175	G77.
	Gly	L∈u	: Glu	Ser 180		Val	Leu	Lys	Tyr 185		Ala	Gln	Asn	Met 19(Glu	Leu
40	Gln	Asr	Lys 195		Gln	Leu	Leu	Glu 200	Glu	Gln	Asn	Leu	Ser 205	Leu	Leu	Ası
	Gln	Leu 210		lys	: Leu	Gln	Ala 215	Met	Val	Ile	Glu	11∈ 220	Ser	Asn	Lys	The
45	Ser 225		ser Ser	Ser	Thr	Cys 23(. Leu	Val	. Leu	Leu 235	Val	. Ser	Ph∈	Cys	Leu 240
50	Leu	: Le	ı Val	Pro	245		Ţуr	· Ser	Ser	250		Arg	, Gly	· Ser	Leu 25!	ı Pro
	Ala	a Gli	His ي	Gly 260		Leu	Sei	r Arg	Glr 265		ı Arç	ı Ala	a Leu	270	Ser	Glv
55	Asp	Pr	275 275		n Leu	ı Glu	l Lei	28(a Leu	ı Glr	sei	285 285	ı Val	Pro	Ly:
	Asj	2 5€: 29		r Hi	e Glr	ı Trp	29!		o Gly	y Sei	Asp	300 300		l Lev	ı Glr	n Ala

60 Pro Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser

```
3 C 5
                     310
                                         315
     Ala Glu Fro Pro Leu Glu Trp Fro Phe Pro Asp Leu Ser Ser Glu Pro
                  325
                         33C 331
     Leu Cys Arg Gly Fro Ile Leu Fro Leu Gln Ala Ash Leu Thr Ard Lys
                              345
     Gly Gly Trp Leu Fro Thr Gly Ser Pro Ser Val Ile Leu Gln Asp Arg
10
      36t 36t
     Tyr Ser Gly
        376
15
     (2) INFORMATION FOR SEQ ID NO: 260:
           (i) SEQUENCE CHARACTERISTICS:
20
                 (A) LENGTH: 12 amino acids
                  (B) TYPE: amino ació
                  (D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26(+
25
     Cys Arg Cys Ala Ser Gly Phe Thr Gly Glu Asp Cys
              ĩ
30
    (2) INFORMATION FOR SEÇ ID NO: 261:
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 11 amino acids
                  (E) TYPE: amino acid
35
                 (D) TOPOLOGY: linear
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 261:
     Cys Thr Cys Gln Val Gly Phe Thr Gly Lys Glu Cys
      40
     (2) INFORMATION FOR SEQ ID NO: 262:
4.5
           (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 12 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 262:
50
     Cys Leu Asn Leu Fro Gly Ser Tyr Gln Cys Gln Cys
                 ŗ
                                   1 (
55
     (2) INFORMATION FOR SEC ID NO: 263:
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 12 amino acids
60
                  (B) TYPE: amino acid
```

```
(D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 263:
     Cys Lys Cys Leu Thr Gly Phe Thr Gly Gln Lys Cys
     (2) INFORMATION FOR SEC ID NO: 264:
10
            (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 12 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 264.
15
     Cys Gln Cys Leu Gln Gly Phe Thr Gly Gln Tyr Cys
      Ž Ę
20
      (2) INFORMATION FOR SEC ID NC: 265:
             (i) SEQUENCE CHARACTERISTICS:
25
                   (A) LENGTH: 127 amino aciás
                    (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 265:
      Gly Leu Ala Cys Trp Leu Ala Gly Val Ile Phe Ile Asp Arg lys Arg
30
                               10
      Thr Gly Asp Ala Ile Ser Val Met Ser Glu Val Ala Gln Thr Leu Leu
35
      Thr Gln Asp Val Xaa Val Trp Val Phe Pro Glu Gly Thr Arg Asn His
      Asr. Gly Ser Met Leu Pro Phe Lys Arg Gly Ala Phe His Leu Ala Val
40
      Gln Ala Gln Val Pro Ile Val Fro Ile Val Met Ser Ser Tyr Gln Asp
      Phe Tyr Cys Lys Lys Glu Arg Arg Phe Thr Ser Gly Gln Cys Gln Val
45
                                          90
      Arg Val Leu Pro Pro Val Frc Thr Glu Gly Leu Thr Pro Asp Asp Val
                        105
 50
      Pro Ala Leu Ala Asp Arg Val Arg His Ser Met Leu His Cys Phe
                                120
              115
 55
       (2) INFORMATION FOR SEQ ID NO: 266:
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 98 amino acids
                     (B) TYPE: amino acid
 60
```

```
(D) TOPOLOGY: linear
             (xi, SEQUENCE DESCRIPTION: SEQ ID NO: 260
      Fro Ser Ala Lys Tyr Phe Phe Lys Met Ala Phe Tyr Ash Gly Trp Ile
                                         1(
      Leu Pre Leu Ala Val Leu Ala Ile Pro Val Cys Ala Val Arg Gly Arg
                 20
                                    25
     Ash Val Glu Ash Met Lys Ile Leu Arg Leu Met Leu Leu His Ile Lys
      Tyr Leu Tyr Gly Ile Arg Val Glu Val Arg Gly Ala His His Phe Pro
                            E E
15
      Fro Ser Gln Fro Tyr Val Val Val Ser Ash His Gln Ser Ser Leu Asr
             7(
      Leu Leu Gly Met Met Glu Val Leu Frc Gly Arg Cys Val Prc Ile Ala
20
                             9(
     Lys Arc
25
     (2) INFORMATION FOR SEQ ID NO: 267:
            (i) SEQUENCE CHARACTERISTICS:
30
                   (A) LENGTH: 9 amino acids
                   (E) TYPE: amino ació
                   (D) TOPCLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEÇ ID NO: 267:
35
     Thr Val Phe Arg Glu Ile Ser Thr Asp
40
     (2) INFORMATICH FOR SEQ ID NO: 268:
            (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 11 amino acids
                   (B) TYPE: amino acid
45
                   (D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 268:
     Lev Trp Ala Gly Ser Ala Gly Trp Pro Ala Gly
       1
50
     (2) INFORMATION FOR SEQ ID NO: 269:
55
            (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 29 amino aciós
                   (B) TYPE: aminc acid
                   (D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 269:
60
```

```
Ser lie Leu Gly Ile Ile Ser Val Fro Leu Ser Ile Gly Tyr Cys Ala
     Ser Lys His Ala Leu Arg Gly Phe Phe Asn Gly Leu Arg
     (2) INFORMATION FOR SEC ID NO: 270:
10
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 8 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 270:
15
     Met Ala Tyr His Gly Leu Thr Val
      1 5
20
     (2) INFORMATION FOR SEQ ID NO: 271:
            (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 6 amino acids
25
                    (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 271:
30
      lie Ser Ala Ala Arg Val
      (2) INFORMATION FOR SEQ ID NO: 272:
35
             (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 11 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
40
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 272:
      Pro Asp Val Ser Glu Phe Met Thr Arg Leu Phe
                  5
45
      (2) INFORMATION FOR SEQ ID NO: 273:
             (i) SEQUENCE CHARACTERISTICS:
 50
                    (A) LENGTH: 17 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 273:
 55
      Phe Asp Pro Val Arg Val Asp Ile Thr Ser Lys Gly Lys Met Arg Ala
           į 10
      Arc
```

300					
Applicants or agents flic reference number	°\$001PC	International application	Unassigned		

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13tis)

A. The indications made below relate to the nacroolganism rete on page 64 line N	ened to in the description.
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution. American Type Culture C	Collection.
Address of depositary institution (including postal code and country) 12301 Parkiawn Drive Rockville, Maryland 20852 United States of America	untry)
Date of deposit February 26, 1991	Accession Number 97901
C. ADDITIONAL INDICATIONS were trans it not oppose	cable: This information is continued on an additional sheet
	IONS ARE MADE (it the indications are not for all designated States)
E. SEFARATE FURNISHING OF INDICATIONS theat The indications listed below will be sufmitted to the International	ave blank it not applicable. All butcau later (specify the general nature of the malcations, e.g., "Accession
Number of Deposit"	
For receiving Office use only	For International Bureau use only
his sheet was received with the international application	This sheet was received by the International Bureau on
Authorized officer	Authorized office

	361		
Applicants or agents file	28001PCT	internationa: aprikatic	Unassigned
reference number		<u> </u>	

(PCT Rule 13his

A. The indications made below relate to the microorganism on page 64	reterred to in the description N/A
B. IDENTIFICATION OF DEPOSIT	Further denosits are identified on an additional sheet
Name of depositary institution American Type Culture	e Collection
Address of depositary institution (including postal code and 12301 Parklawn Drive Rockville, Maryland 20852	Country)
United States of America	
Date of deposit February 26, 1997	Accession Number 97898
C. ADDITIONAL INDICATIONS HELVE LIGHA II NOI OF	of ucable. This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDICA	STIONS ARE MADE (1) the inaications are not for all designated States)
E. SEPARATE FURNISHING OF INDICATIONS	(leave blank ij noi applicable
The indications listed below will be submitted to the internal Number of Deposit";	tional Bureau later (specify ine general nature of the indications, e.g., Accessio
For receiving Office use only	For international Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau or
Authorized office	Authorized officer

		362		
Applicant's or agent's file	25001PC"	international aprilication	Unassigner	
reference number				

(PCT Rule 13his

on page 64 line	N/A
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution American Type Cult	ure Collection
Address of depositary institution (including postal code of	na country
12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209044
C. ADDITIONAL INDICATIONS (HEAVE CLOTH A) THOU	application. This information is continued on an additional sheet.
D. DESIGNATED STATES FOR WHICH INDIC	CATIONS ARE MADE (i) the indications are not for all designated States)
. SEPARATE FURNISHING OF INDICATION	
The indications listed below will be submitted to the inten- Number of Deposit")	national Buteau later ispecify the general nature of the maicallons, e.g., 'Accession (x,y)
Constant Office and Only	Control Durage use only
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on
Authorized officer	Authorized office:
1 k />	

Form PCT/RO/134 (July 1992)

 $\S^{\bullet}(C) = \{ (1, 2, \dots, n) \mid \text{ as value}(\mathcal{L}) \}$

Form PCT/RO/134 (July 1992

363

	202			
		1	U-seeigmad	
Applicants or agent's file	⊇\$00 1PC ".	International application	Unassigned	
reference number				

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis

A. The indications made below relate to the microorganism re on page 64 . line	eterred to in the description. N/A
B. IDENTIFICATION OF DEPOSIT	i urther deposits are identified on an additional sheet
Name of depositary institution American Type Culture	Collection
Address of depositary institution (including postal code and code) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	ountry)
Date of deposit February 26, 1997	Accession Number 97899
C. ADDITIONAL INDICATIONS THEOre trans it not app	cucable. This information is continued on an accitional sheet
D. DESIGNATED STATES FOR WHICH INDICA	TIONS ARE MADE (if the indications are not for all designated States)
E. SEPARATE FURNISHING OF INDICATIONS	(leave blank if not applicable)
The indications listed below will be submitted to the international Number of Deposit")	tional Buteau latet (specify the general nature of the indications, e.g., "Accession
For receiving Office use only	For international Bureau use only
This sheet was received with the international application. Authorized officen	This sheet was received by the International Eureau on

Applicants or acents to	08001PC1	international april	ucatic Unassigne	c

(PCT Rule 13bis

A. The indications made below relate to to on page $-\frac{\epsilon}{2}$	the microorganism referred to time	c in the description.
B. IDENTIFICATION OF DEPOSI	דו	Further deposits are identified on ar. additional sheet
Name of depositary institution Amer	ncan Type Culture Collec	tior.
Address of depositary institution circlian	ng postal coae and country	
12301 Parkiawn Drive Rockville, Maryland 20852 United States of America		
Date of deposit May 15, 1997	A	ccession Number 209045
C. ADDITIONAL INDICATIONS	Heave trank II noi applicable.	This information is continued on an additional sheet
D. DESIGNATED STATES FOR V		ARE MADL (if the inaications are not for all aesignated States)
E. SEPARATE FURNISHING OF	INDICATIONS (wave bid	ink II noi appucable
The indications listed below will be subminumber of Deposit":	ined to the international bu	reau tatet ispecin ine general nature of the indications, e.g., 'Accession
l or receiving Office	use only	For international bureau use only
This sheet was received with the interr	эгислаг аррисвиот	This sheet was received by the international Eureau on
Authorized officer		Authorized office:

	36	5		
Applicants or agents fill	2S001PCT	international applicatio.	Unassigned	
reterence number				

(PCT Rule 13bis)

A. The indications made below relate to the microorganism on page 64 , line	reterred to in the description. N/A
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution American Type Cultur	re Collection
Address of depositary institution finctuaing postal code and 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	: country;
Date of deposit February 26, 1997	Accession Number 97900
C. ADDITIONAL INDICATIONS HEAVE BIGHA II HOLD	princation. This information is continued on an additional sheet.
D. DESIGNATED STATES FOR WHICH INDIC.	ATIONS ARE MADE (if the indications are not for all designated States)
E. SEPARATE FURNISHING OF INDICATIONS	S (leave blank ij noi applicable)
The indications listed below will be submitted to the international Number of Deposit"	ational Bureau later (specify the general nature of the indications, e.g., "Accessi
For receiving Office use only	For international Bureau use only
This sneer was received with the international application	This sheet was received by the International Bureau on
Authorized office:	Authorized officer

Form PCT/RO/134 (July 1992)

Grande Land

	366			
Applicants or agents the	5001PC"	international applicatio	Unassigned	- <u>-</u> 1, <u>1</u>
reterence number				

(PCT Rule 13bis)

A. The indications made below relate to the fluctoors on page 6.	time N/A
B. IDENTIFICATION OF DEPOSIT	further deposits are identified on an additional sneet in
Name of depositary institution. American Type	Culture Collection.
Address of depositors institution (including postal coll 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	ige and country:
Date of deposit May 15, 1997	Accession Number 20904r
C. ADDITIONAL INDICATIONS THEORE CHAPK	v.nc. applicable) This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH IN	NDICATIONS ARE MADE (1) the inaccations are not for all aesignated States)
E. SEFARATE FURNISHING OF INDICAT	HONS (seave blank if not applicable)
The indications listed below will be submitted to the Number of Deposit"	International Bureau later (specify the general nature of the indications, e.g., "Accession
For receiving Office use only	For international Bureau use only
Phis sheet was received with the international appli	This sheet was received by the International Bureau on
Authorized officer	Authorized office

		367		
Applicants or agents file	'S001PC'	internationa, applicane	Unassignen	
reference number				

(PCT Rule 13tis

3. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
name of depositary institution. American Type Cultu-	
Address of depositary institution (<i>including postal code an</i> 12301 Parklawn Drive Rockville, Maryland, 20852 United States of America	ag country)
Date of deposit April 28, 1997	Accession Number 209010
C. ADDITIONAL INDICATIONS HELVE LIGHE II FIGH	appucable) This information is continued on an additional sheet
	CATIONS ADE MADE in the indications are not for all designated States)
D. DESIGNATED STATES FOR WHICH INDIC	CATIONS ARE MADE (1) the inaccations are not for all designated States,
	CATIONS ARE MADE (1) the inaccations are not for all designated States,
E. SEPARATE FURNISHING OF INDICATION The indications listed below will be submitted to the Inter	NS (ieave blank i) noi applicable
E. SEPARATE FURNISHING OF INDICATION The indications listed below will be submitted to the Inter- Number of Deposit")	NS lieave blank il noi applicable Tallonal Bureau lalei (specify ine general nature of the indications, e.g., "Access
E. SEPARATE FURNISHING OF INDICATION	NS (leave blank if not applicable Tational Bureau later (specify the general nature of the indications, e.g., 'Access For international Bureau use only

		368			
Applicants or agents the	15001PC		international applicatio	Unassigned	 :
reterence number					

(PCT Rule 13th)

B. IDENTIFICATION OF DEPOSIT	i unities deposits are identified on an additional sheet
Name of depositary institution American Type Culture	e Collection
Address of depositary institution (including postal code and	Country
12301 Parklawn Drive Rockville, Marvlano, 20852 United States of America	
Date of deposit May 29, 19 ^C	Accession Number 209085
C. ADDITIONAL INDICATIONS (wasse trank it not at	This information is continued on an additional sheet
TO THE LOD WHICH INDIC	
D. DESIGNATED STATES FOR WHICH INDICA	ATIONS ARL MADE til the inaications are not for all aesignatea States)
E. SEFARATE FURNISHING OF INDICATIONS	S Heave blank II nol applicable
E. SEFARATE FURNISHING OF INDICATIONS The indications listed below will be submitted to the internal	S lieave tians il noi attiikabik
E. SEFARATE FURNISHING OF INDICATIONS The incications listed below will be surmitted to the internal Number of Deposit ²⁵	S (Leave blank il noi applicable ational Buteau latet (specify) the general nature of the indications, e.g., Accessio

	36	è		
Applicants or agents the	'S001PC'.	internationa, application	Unassigneo	
reference number		l		

(PCT Rule 13tus

A. The indications made below relate to the microorganism rete	rrec to in the description.
on page 65	Further deposits are identified on an additional sheet
Name of depositary institution American Type Culture Co	ollection
Address of depositary institution (including postal code and could 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	ntry:
Date of deposit February 26, 1997	Accession Number 97897
C. ADDITIONAL INDICATIONS ILLANG BIGHA II NOI APPILL	This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDICATI	ONS ARE MADE (1) the indications are not for all designated States)
E. SEPARATE FURNISHING OF INDICATIONS (ied	ave blank u noi appucable
The indications listed below will be submitted to the internation Number of Deposit")	nal Bureau later (specify the general nature of the indications, e.g., 'Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau or.
Authorized officer	Authorized office:

Form PCT/RO/134 (Jbly 1992)

s - + 11.11 F

		370		
Applicant's or agent's 11. reference number	S001PCT	international application	Unassigner	<u></u>

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microbion page $-\epsilon^{\frac{1}{2}}$	reathsm referred to in the description. . The N/A
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution American Type	e Culture Collection
Address of depositant institution (including postal collision) 12301 Parkiawn Drive Rockville, Marviand 20851 United States of America	oge and country
Date of deposit May 15, 1997	Accession Number 20904?
C. ADDITIONAL INDICATIONS HEAVE EIGHT	This information is continued on an additional sheet
E. SEPARATE FURNISHING OF INDICAT	NDICATIONS ARE MADE up the indications are not for all designated States)
	International Eureau later ispecify the general nature of the malcations, e.g., 'Accession
	•
For receiving Office use only.	For International Bureau use only This sheet was received by the international Bureau on
Authorized officer	Authorized officer

		371			
Applicants or agents fill	S001PCT	internationa:	application	Unassigner	
reference number		<u> </u>			

(PCT kule 13bis)

A. The indications made below relate to the microorgal on page [nism referred to in the description
. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet in
Name of depositant institution American Type C	ulture Collection
Address of acpositary institution (including postal code) 12301 Parklawn Drive Rockville, Marviand, 20852 United States of America	ang country
Date of deposit September 4, 1997	Accession Number 20923¢
C. ADDITIONAL INDICATIONS (wave blank it	no applicable. This information is continued on an additional sneet
E. SEPARATE FURNISHING OF INDICATI	ONS (seave blank ij noj applicable)
	nternational Burcau later (specify the general nature of the indications, e.g., "Acces
For receiving Office use only	For International Bureau use only
This sheet was received with the international applica	This sheet was received by the International Bureau on
Authorized office:	: Authorized office:

Form PCT/RO/134 Nuly 1992

		372		
Applicants of acents to	SOULPC':	international application	Onassigned	* *** 45 15
reference numbe		1		

(PCT Rule 13tis)

A. The indications made peach relate to the microorganism referred to in the description on page 5.5 line N/A				
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet.			
name of depositors in stitution. American Ty	ype Culture Collection			
Address of depositary institution <i>(including posic</i> 12301 Parkiawn Drive Rockville, Marviand 20852 United States of America	a. cour una country			
Date of deposit May 29, 1997	Accession Number 2090&			
C. ADDITIONAL INDICATIONS most t	This information is continued on an additional sheet			
). DESIGNATED STATES FOR WHICH	H INDICATIONS ARE MADE, in the indications are not for all designated States)			
E. SEFARATE FURNISHING OF INDIC The incications listed below will be surmitted to Number of Deposit"	CATIONS (leave blank it not applicable) the international Bureau later (specify the general nature of the indications, e.g., "Accessing the content of the indications, e.g.,"			
For receiving Office use on	For international Bureau use only			
For receiving Office use on				
The sheet was received with the international	application. This sheet was received by the International Bureau on			

			_	_
Applicants	or	acent	5	Tuc
reference ni	um	DC.		

5001PCT

international application

Unassignec

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism on page 7t	n reterred to in the description. N/A
3. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution. American Type Culti	ure Collection
Address of depositary institution (including postal code an 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	ac country)
Date of deposit May 15, 1997	Accession Number 209048
C. ADDITIONAL INDICATIONS (MOVE FIGHA II FIG.	approcable. This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDIC	CATIONS ARE MADE (i) the inaccasions are not for all aesignated States)
E. SEFARATE FURNISHING OF INDICATION	NS (leave blank if not applicable)
The indications listed below will be submitted to the Inter Number of Deposit")	mational Bureau later (specify the general nature of the indications, e.g., "Accession
For receiving Office use only	For international Bureau use only
This sheet was received with the international application	
Authorized officer	Authorized officer

Form. PCT/RO/134 (July 1992)

La Carrier Company

Appricants or agents file	SUCIPC	international aprincation	Unassigned	1
reference number		1	Cita Signed	
reference numbe				

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis.

A. The indications made below relate to the microorganism on page 7t line	referred to in the descriptio: N/A
B. IDENTIFICATION OF DLPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution. American Type Culture	e Collection
Address of depositary institution (including postal code one of	country)
12301 Parklawn Drive Rock ville, Marvland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97902
C. ADDITIONAL INDICATIONS more trans tract app	rucabie. This information is continued on an additional sheet.
D. DESIGNATED STATES FOR WHICH INDICA	TIONS ARE MADL (1) the indicotions are not for all designated States)
E. SEFARATE FURNISHING OF INDICATIONS	<u> </u>
The incications usted below will be submitted to the internati Number of Deposit")	ional Buteau later ispecini ine general nature of the indications, e.g., 'Accession
For receiving Office use only	For international Bureau use only
This sheet was received with the international application.	This sheet was received by the International Bureau or
Authorized officer	Authorized office:

	37.	5		
Apriicants or agents the	S001PCC	International ar piccation	Unassigned	The same of the sa
reference number				

(PCT Kule 13tis

A. The indications made below relate to the microorganism referred to in the description on page 77				
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet			
Name of depositary institution American Type Culture Coll	ectior.			
Address of depositary institution (including postal code and country) 12301 Parkiawn Drive Rockville, Maryland 20852 United States of America	ועך			
Date of deposit February 26, 1997	Accession Number 97903			
C. ADDITIONAL INDICATIONS HEAVE CLOTH II THE OFFICED	This information is continued on an additional shee:			
D. DESIGNATED STATES FOR WHICH INDICATION				
E. SEPARATE FURNISHING OF INDICATIONS neave The indications listed below will be submitted to the international Number of Deposit")	blank if not applicable: Bureau later (specify the general nature of the indications, e.g., "Accession			
For receiving Office use only	For International Bureau use only			
This sheet was received with the international application.	This sheet was received by the International Bureau on			
Authorized officer	Authorized office			

Applicants of agents fill reference number

S001PC7 International application

Upassigned

The second second

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Kuie 13bis)

A. The indications made telew relate to the microorganism reterred to in the description on page				
B. IDENTIFICATION OF DEPOSIT	Eurther deposits are identified of, ar, additional sheet			
Name of depositant institution American Type Culture (C ellection			
Address of depositary institution cincinaing postal coae and coae and coae are coae and coae are coae	unir			
Date of deposit May 15, 1997	Accession Number 209049			
C. ADDITIONAL INDICATIONS THEORY FLORE IT INC. OFF	icatu. This information is continued on an additional sheet.			
D. DESIGNATED STATES FOR WHICH INDICAT	IONS ARE MADE (if the inaications are not for all aesignated States)			
E. SEPARATE FURNISHING OF INDICATIONS (ii				
The indications listed below will be submitted to the internation Number of Deposit")	mal bureau later ispecify the general nature of the maicallons, e.g., "Accession			
For receiving Office use only	For International Bureau use only			
This sheet was received with the international application	This sheet was received by the International Bureau on			
Authorized officer	Authorized officer			

Form PCT/RO/134 (July 1992)

assume the second

	311			
Applicants or agents file reference number	S001PC1	international application	Unassigned	;=====================================
(leference name		•		

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis.

on page 8(, line	N/A
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution. American Type Culture	e Collectior.
Address of depositary institution (including postal code and 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	country
Date of deposit February 26, 1997	Accession Number 97904
C. ADDITIONAL INDICATIONS (seeve biank if not ap	princation. This information is continued on an additional sheet.
D. DESIGNATED STATES FOR WHICH INDICA	ATIONS ARE MADE (if the inaications are not for all designated States)
E. SEPARATE FURNISHING OF INDICATIONS	
The indications listed below will be submitted to the internal Number of Deposit")	itional bureau later ispecify the general nature of the malcations, e.g., "Accession of the malcations
For receiving Office use only	For International Bureau use only

Applicants or agents for	S0C1PC7.	inteniational application	Unassign ec	1
reference number				

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

on page 80	ate to the microorganism reten jine N/A	
B. IDENTIFICATION OF DE	POSIT	Further deposits are identified or, an additional sheet
Name of depositary institution	American Type Culture Co	llection
Address of depositary institution (in 1230) Parkiawn Drive	ciuaing postai coae ana coun	ry)
Rockville, Marviand 20852 United States of America		
Date of deposit May 15, 1997		Accession Number 20905(
C. ADDITIONAL INDICATI	ONS theave trank it not applica-	This information is continued on an additional sheet
D. DESIGNATED STATES FO	OR WHICH INDICATIO	NS ARE MADE (if the inaccations are not for all designated States)
E. SEPARATE FURNISHING	OF INDICATIONS (leav	e blank if not applicable)
		Bureau later (specify the general nature of the indications, e.g., "Accessi
Number of Deposit")	Submitted to the international	Sales later (speedy) the general harder of the
For receiving O	ffice use only	For International Bureau use only
For receiving O		For International Bureau use only This sheet was received by the International Bureau on:

Applicants or a	gents that
reference numb-	c ·

5001PCT

international approcise-

l-nassigned

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis.

. The indications made below relate to the filleroorganis on page 87 . line	: N/A
IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
ame of depositary institution. American Type Cult	cure Collection
ddress of depositary institution (including postal code of 2301 Parklawn Drive ockville, Maryland 20852 United States of America	na country)
ate of deposit April 4, 1997	Accession Number 9797t
. ADDITIONAL INDICATIONS Heave tians is no	This information is continued on an additional sheet
DESIGNATED STATES FOR WHICH INDI	CATIONS ARE MADE (if the inaccations are not for all aesignated States)
. SEPARATE FURNISHING OF INDICATIO	NS (leave blank ij noi applicable)
The indications listed below will be submitted to the inte Number of Deposit".	Thational bureau later (specify the general nature of the maications, e.g., Access
For receiving Office use only	For International Bureau use only
this sheet was received with the international application	This sheet was received by the international Eureau on
Authorized office:	Authorized officer

Form. PCT/kO/134 (July 1992,

Sept. September

Applicants or agents file	5061PC%	International eppircane:	Unassigned	7 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 -
reference number				

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis.

on page	ns made below relate to the micro	organism referred to in the description. . line N/A
3. IDENTIFIC	CATION OF DEPOSIT	Further decesses are identified on an additional sheet
Name of deposita		pe Culture Collection
Address of depos	mary institution (including postal	coas ana country)
12301 Parklawi Rockville, Mar United States o	yland 20852	
Date of deposit	May 15, 1997	Accession Number 209047
C. ADDITION	NAL INDICATIONS (leave bia	ink it not applicable) This information is continued on an additional sheet
DESIGNAT	TO STATES FOR WHICH	INDICATIONS ARE MADE (if the indications are not for all designated States)
. DESIGNAT	EDSTATES FOR WINCH	INDICATIONS ARE MADE (if the indications are not for all designated states)
. SEPARATE	FURNISHING OF INDICA	ATIONS (leave blank if not applicable)
The indications in	sted below will be submitted to th	e international Bureau later (specify the general nature of the indications, e.g., "Accession
, ,	,	
	For receiving Office use only	For International Bureau use only
	as received with the international app	
	•	
Authorized officer		

15

20

What Is Claimed Is:

- 1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:
- (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X:
- (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z. which is hybridizable to SEQ ID NO:X:
 - (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z which is hybridizable to SEQ ID NO:X:
 - (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X, having biological activity:
 - (f) a polynucleotide which is a variant of SEQ ID NO:X:
 - (ε) a polynucleotide which is an allelic variant of SEQ ID NO:X:
 - (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y:
- (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.
- The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.
- The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

4. The isolated nucleic acid molecule of claim I, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

5

- 5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.
- 10 6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.
- 7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.
 - 8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.
- 20 9. A recombinant host cell produced by the method of claim 8.
 - 10. The recombinant host cell of claim 9 comprising vector sequences.
- 11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:
 - (a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z:
 - (b) a polypeptide tragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;
- 30 (c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z:
 - (d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z:
- (e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z:
 - (f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z:

20

30

U1 - 114 A

- (g) a variant of SEQ ID NO:Y:
- (h) an allelic variant of SEQ ID NO:Y; or
- (i) a species homologue of the SEQ ID NO:Y.
- The isolated polypeptide of claim 11, wherein the secreted form or the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.
 - 13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.
- 14. A recombinant host cell that expresses the isolated polypeptide of claim 11.
 - 15. A method of making an isolated polypeptide comprising:
- 15 (a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and
 - (b) recovering said polypeptide.
 - 16. The polypeptide produced by claim 15.
 - 17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 11 or the polynucleotide of claim 1.
- 25 18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:
 - (a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and
 - (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.
 - 19. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:
- (a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and
 - (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

- 20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:
 - (a) contacting the polypeptide of claim 11 with a binding partner; and
- 5 (b) determining whether the binding partner effects an activity of the polypeptide.
 - 21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.
- 10 22. A method of identifying an activity in a biological assay, wherein the method comprises:
 - (a) expressing SEQ ID NO:X in a cell;
 - (b) isolating the supernatant;
 - (c) detecting an activity in a biological assay; and
- 15 (d) identifying the protein in the supernatant having the activity.
 - 23. The product produced by the method of claim 22.

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referr on page 64 , line N/A	
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet 🔀
Name of depositary institution American Type Culture Col	lection
Address of depositary institution (including postal code and count 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	rry)
Date of deposit February 26, 1997	Accession Number 97900
C. ADDITIONAL INDICATIONS (leave blank if not applical	ble) This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDICATIO	ONS ARE MADE (if the indications are not for all designated States)
E. SEPARATE FURNISHING OF INDICATIONS (leave	e blank if not applicable) Bureau later (specify the general nature of the indications, e.g., "Accession
Number of Deposit")	
For receiving Office use only This sheet was received with the international application	This sheet was received by the International Bureau on
Authorized officer	Authorized officer

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

Page 2

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the applicant is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicants or agents file	PS001PC"	International application	Unassigned	
		111101110111111111111111111111111111111		
reterence number		1		

(PCT Rule 13bis)

A. The indications made below relate to the microorganism refer on page $\frac{65}{100}$, line $\frac{N}{2}$	
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution — American Type Culture Co	ollection
Address of depositary institution (including postal code and cour 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	ntry)
Date of deposit May 15, 1997	Accession Number 209043
C. ADDITIONAL INDICATIONS (seave blank if not applice	able) This information is continued on an additional sheet
made available until the publication of the mention of the gapplication has been refused or withdrawn or is deemed to nominated by the person requesting the sample (Rule 28 (4	be withdrawn, only by the issue of such a sample to an expert
E. SEPARATE FURNISHING OF INDICATIONS (lea The indications listed below will be submitted to the International Number of Deposit")	ave blank if not applicable) al Burcau later (specify the general nature of the indications, e.g., "Accessio
For receiving Office use only This sheet was received with the international application	For International Bureau use only This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

Page 2

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

(PCT Rule 13bis)

. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
ame of depositary institution American Type Culture C	Collection
address of depositary institution (including postal code ana cou	untry)
2301 Parklawn Drive Rockville, Maryland 20852 Jnited States of America	
vate of deposit May 15, 1997	Accession Number 209044
C. ADDITIONAL INDICATIONS (leave blank if not appli	cable) This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDICAT	IONS ARE MADE (if the indications are not for all designated States)
D. DESIGNATED STATES FOR WHICH INDICAT	IONS ARE MADE (if the indications are not for all designated States)
E SEPARATE FURNISHING OF INDICATIONS (I	eave blank if not applicable)
E. SEPARATE FURNISHING OF INDICATIONS (International Internations listed below will be submitted to the Internation	eave blank if not applicable)
E. SEPARATE FURNISHING OF INDICATIONS (International Internations listed below will be submitted to the Internation	eave blank if not applicable)
E. SEPARATE FURNISHING OF INDICATIONS (to The indications listed below will be submitted to the Internation Number of Deposit")	
E SEPARATE FURNISHING OF INDICATIONS (I	eave blank if not applicable) nal Bureau later (specify the general nature of the indications, e.g., "Access

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

PARTY (. W) GENERAL E

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the applicant is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism refe on page 65 , line N/	
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet 🖂
Name of depositary institution American Type Culture Co	ollection
Address of depositary institution (including postal code and could 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	ntry)
Date of deposit May 15, 1997	Accession Number 209045 «
C. ADDITIONAL INDICATIONS (leave blank if not applic	cable) This information is continued on an additional sheet
made available until the publication of the mention of the gapplication has been refused or withdrawn or is deemed to nominated by the person requesting the sample (Rule 28 (4	be withdrawn, only by the issue of such a sample to all expert
·	
E. SEPARATE FURNISHING OF INDICATIONS (lea	
The indications listed below will be submitted to the International Number of Deposit")	al Bureau later (specify the general nature of the indications, e.g., "Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

Applicant's or agent's file PS001PCT	International application 10.	Unassigned	to the transmittee of the
reference number	l		a despite en en constitución de la constitución de

A. The indications made below relate to the microorganism on page 64 , line	
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet 🔀
Name of depositary institution American Type Cult	ture Collection
Address of depositary institution (including postal code and 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	nd country)
Date of deposit May 15, 1997	Accession Number 209046
C. ADDITIONAL INDICATIONS (leave blank if not	applicable) This information is continued on an additional sheet
application has been refused or withdrawn or is deen nominated by the person requesting the sample (Rule	of the grant of the European patent or until the date on which ned to be withdrawn, only by the issue of such a sample to an expert e 28 (4) EPC). CATIONS ARE MADE (if the indications are not for all designated States)
E. SEPARATE FURNISHING OF INDICATION	NS (leave blank if not applicable)
The indications listed below will be submitted to the Inter Number of Deposit")	mational Bureau later (specify the general nature of the indications, e.g., "Accessio
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	on This sheet was received by the International Bureau on.
Authorized officer	Authorized officer

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

A. The indications made below relate to the microorganism ref	ferred to in the description ./A
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet 🖂
Name of depositary institution American Type Culture C	Collection
Address of depositary institution (including postal code and con 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	untry)
Date of deposit May 15, 1997	Accession Number 209047
C. ADDITIONAL INDICATIONS (leave blank if not applied	cable) This information is continued on an additional sheet
made available until the publication of the mention of the	be withdrawn, only by the issue of such a sample to an expert
D. DESIGNATED STATES FOR WHICH INDICATI	ONS ARE MADE (if the indications are not for all designated States)
E. SEPARATE FURNISHING OF INDICATIONS (lea	zve blank ij not applicable)
The indications listed below will be submitted to the Internation. Number of Deposit")	al Bureau later (specify the general nature of the indications, e.g., "Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

A. The indications made below relate to the microorganism refe on page -76 , line $-N/$	
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet 🔀
Name of depositary institution American Type Culture Co	ollection
Address of depositary institution (including postal code and counting 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	ntry)
Date of deposit May 15, 1997	Accession Number 209048
C. ADDITIONAL INDICATIONS (leave blank if not applic	rable) This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDICATION	be Withdrawn, only by the issue of such a sumple to all original
- ···	
E. SEPARATE FURNISHING OF INDICATIONS (lea	zve blank if not applicable)
The indications listed below will be submitted to the Internations Number of Deposit')	al Bureau later (specify the general nature of the indications, e.g., "Accessio
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the applicant is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

Applicant's or agent's file	PS001PCT	international application	.,1	Unassigned		
reterence number					4	

(PCT Rule 13bis)

A. The indications made below relate to the microorganism reference on page 77 , line N/A	
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution American Type Culture Co	ollection
Address of depositary institution (including postal code and coun 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	ıtry)
Date of deposit May 15, 1997	Accession Number 209049
C. ADDITIONAL INDICATIONS (leave blank if not applica	able) This information is continued on an additional sheet
nominated by the person requesting the sample (Rule 28 (4	be withdrawn, only by the issue of such a sample to an expert
TIONS .	
E. SEPARATE FURNISHING OF INDICATIONS (lear The indications listed below will be submitted to the International Number of Deposit")	al Bureau later (specify the general nature of the indications, e.g., "Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

service A.

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

		1		Unaccionad	
Applicants or agent's file	□ 05001PCT	International application	,	Unassignec	
reference number				4.00 m	The state of the s
(Citicination)					

A. The indications made below relate to the microorganism ret	ferred to in the description
on page 80 .line N	N/A
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet 🔀
Name of depositary institution American Type Culture C	Collection
Address of depositary institution (including postal code and co	ountry)
12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
	Accession Number 209050 "
Date of deposit May 15, 1997	Accession rumoer 207070 "
C. ADDITIONAL INDICATIONS (leave blank if not appl	licable) This information is continued on an additional sheet
made available until the publication of the mention of the application has been refused or withdrawn or is deemed t nominated by the person requesting the sample (Rule 28	to be withdrawn, only by the issue of the
E. SEPARATE FURNISHING OF INDICATIONS	leave blank if not applicable)
The indications listed below will be submitted to the Internation Number of Deposit')	onal Bureau later (specify the general nature of the indications, e.g., "Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on
Authorized officer	Authorized officer

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

Applicants or agents file - eterence number	PS001PCT	International application	ю.	unassigned	

(PCT Rule 13bis)

A. The indications made below relate to the microorganism refer on page 73 , line N/A	•
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution American Type Culture Co	llection
Address of depositary institution (including postal code and count	וניזו
12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit September 4, 1997	Accession Number 209236 .
C. ADDITIONAL INDICATIONS (leave blank if not applicable)	This information is continued on an additional sheet
In respect to those designations in which a European Patent in made available until the publication of the mention of the grapplication has been refused or withdrawn or is deemed to be nominated by the person requesting the sample (Rule 28 (4)) D. DESIGNATED STATES FOR WHICH INDICATION	ant of the European patent or until the date on which e withdrawn, only by the issue of such a sample to an expert EPC).
E. SEPARATE FURNISHING OF INDICATIONS (leave	
The indications listed below will be submitted to the International Internation	Bureau later (specify the general nature of the indications, e.g., "Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

Form PCT:RO/134l(July 1992)

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

International application No. Unassigned

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

A. The indications made below relate to the microorganism referr on page 65 line N/A	
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution American Type Culture Col	llection
Address of depositary institution (including postal code ana count 12301 Parklawn Drive	rry)
Rockville, Maryland 20852 United States of America	
Date of deposit April 28, 1997	Accession Number 209010
C. ADDITIONAL INDICATIONS (leave blank if not applicable	ble) This information is continued on an additional sheet
In respect to those designations in which a European Patent made available until the publication of the mention of the grapplication has been refused or withdrawn or is deemed to be nominated by the person requesting the sample (Rule 28 (4)) D. DESIGNATED STATES FOR WHICH INDICATIO	ant of the European patent or until the date on which he withdrawn, only by the issue of such a sample to an expert () EPC).
E. SEPARATE FURNISHING OF INDICATIONS (leave	
The indications listed below will be submitted to the International Number of Deposit")	Bureau later (specify the general nature of the indications, e.g., "Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on
Authorized officer	Authorized officer

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

ENCLOSE WO GENERAL

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the applicant is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

Applicant's or agent's file	PS001PCT	International application	`'0.	Unassigned	
reterence number					to the second of the second

A. The indications made below relate to the microorganism referr on page 65 , line N/A	·
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet 🛛
Name of depositary institution American Type Culture Col	lection
Address of depositary institution (including postal code and count 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	77)
Date of deposit May 29, 1997	Accession Number 209085
C. ADDITIONAL INDICATIONS (leave blank if not applicab	the) This information is continued on an additional sheet
In respect to those designations in which a European Patent is made available until the publication of the mention of the grapplication has been refused or withdrawn or is deemed to be nominated by the person requesting the sample (Rule 28 (4)) D. DESIGNATED STATES FOR WHICH INDICATION	ant of the European patent or until the date on which e withdrawn, only by the issue of such a sample to an expert EPC).
E. SEPARATE FURNISHING OF INDICATIONS (leave	·
Number of Deposit")	Bureau later (specify the general nature of the indications, e.g., "Accession
For receiving Office use only This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

A. The indications made below relate to the microorganism referred to in the description on page 64 , line N/A		
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet 🔀	
Name of depositary institution American Type Culture Co	ollection	
Address of depositary institution (including postal code and county) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	ntry)	
Date of deposit February 26, 1997	Accession Number 97901 .	
C. ADDITIONAL INDICATIONS (leave blank if not applice	able) This information is continued on an additional sheet	
nominated by the person requesting the sample (Rule 26 (be withdrawn, only by the	
E. SEPARATE FURNISHING OF INDICATIONS (lea	ave blank if not applicable)	
The indications listed below will be submitted to the International Number of Deposit")	al Bureau later (specify the general nature of the indications, e.g., "Accession	
For receiving Office use only	This sheet was received by the International Bureau on: Authorized officer	
Authorized officer		

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

The second secon

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

Applicants or agent's file PS001PC1	International application	o.	Unassigned		
reference number				The second of the second	

A. The indications made below relate to the microorganism referred to in the description on page 77 . line N/A .		
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet \boxtimes	
Name of depositary institution American Type Culture Col	llection	
Address of depositary institution (including postal code and count 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	try)	
Date of deposit February 26, 1997	Accession Number 97903	
C. ADDITIONAL INDICATIONS (leave blank if not applicable	ble) This information is continued on an additional sheet	
D. DESIGNATED STATES FOR WHICH INDICATIO	NS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave	e blank if not applicable)	
	Bureau later (specify the general nature of the indications, e.g., "Accessio	
For receiving Office use only	For International Bureau use only	
This sheet was received with the international application	This sheet was received by the International Bureau on:	
Authorized officer	Authorized officer	

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

Applicant's or agent's file	PS001PC1	International application	o. Unassigned	
reference number				

A. The indications made below relate to the microorganism referred to in the description on page 64 . line N/A		
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet 🔀	
Name of depositary institution American Type Culture Co	ollection	
Address of depositary institution (including postal code and county) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	ntry)	
Date of deposit February 26, 1997	Accession Number 97898	
C. ADDITIONAL INDICATIONS (leave blank it not applied	rable) This information is continued on an additional sheet	
made available until the publication of the mention of the application has been refused or withdrawn or is deemed to nominated by the person requesting the sample (Rule 28 (D. DESIGNATED STATES FOR WHICH INDICATION)	ONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (lea		
The indications listed below will be submitted to the Internations Number of Deposit")	al Bureau later (specify the general nature of the indications, e.g., "Accession	
For receiving Office use only	For International Bureau use only	
This sheet was received with the international application	This sheet was received by the International Bureau on	
Authorized officer	Authorized officer	

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the applicant is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

Applicant's or agent's file	PS001PC1	International application No. Unassigned
reterence number		the state of the s

A. The indications made below relate to the microorganism referred to in the description on page 80 , line N/A		
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet.	
Name of depositary institution American Type Culture Co	llection	
Address of depositary institution (including postal code and coun 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	ury)	
Date of deposit February 26, 199	Accession Number 97904	
C. ADDITIONAL INDICATIONS (leave blank if not application)	tible) This information is continued on an additional sheet	
nominated by the person requesting the sample (Rule 28 (4	rant of the European patent or until the date on which be withdrawn, only by the lissue of such a sample to an expert	
E. SEPARATE FURNISHING OF INDICATIONS (leave The indications listed below will be submitted to the International Number of Deposit")	re blank if noi applicable) Buteau later (specify the general nature of the indications, e.g., "Accession	
For receiving Office use only	For International Bureau use only	
This sheet was received with the international application	This sheet was received by the International Bureau on:	
Authorized officer	Authorized officer	

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism reference on page $\frac{73}{100}$, line $\frac{N}{100}$	erred to in the description I/A
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet 🔀
Name of depositary institution American Type Culture C	Collection
Address of depositary institution (<i>including postal code and con</i> 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	untry)
Date of deposit May 29, 1997	Accession Number 209084.
C. ADDITIONAL INDICATIONS (leave blank if not appli	icable) This information is continued on an additional sheet
nominated by the person requesting the sample (Rule 28	o be willidiawii, only by the issue of seen a semper
E. SEPARATE FURNISHING OF INDICATIONS (le The indications listed below will be submitted to the Internation Number of Deposit")	eave blank if not applicable) nal Bureau later (specify the general nature of the indications, e.g., "Access
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

up authority At

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the applicant is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

	100 17 1 to 225	1		7		
Applicants or agents file	PS001PCT	International application	C	Chassigned		
reterence number		i -			1.00	

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism reter on page 64 , line N/A	
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet 🔀
Name of depositary institution American Type Culture Co	ollection
Address of depositary institution (including postal code and coun 12301 Parklawn Drive	ntry)
Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97899
C. ADDITIONAL INDICATIONS (leave blank if not applica	This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDICATION	ONS ARE MADE (if the indications are not for all designated States)
E. SEPARATE FURNISHING OF INDICATIONS (leave	while the applicable
	Il Bureau later (specify the general nature of the indications, e.g., "Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The second second second

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

International application of Unassigned

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 65 line N/A							
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet 🔀						
Name of depositary institution American Type Culture Co	ollection						
Address of depositary institution (including postal code and coun	ury)						
12301 Parklawn Drive Rockville, Maryland 20852 United States of America							
Date of deposit February 26, 1997	Accession Number 97897						
C. ADDITIONAL INDICATIONS (leave blank if not applica	This information is continued on an additional sheet						
In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC). D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)							
E. SEPARATE FURNISHING OF INDICATIONS (leav							
The indications listed below will be submitted to the International Number of Deposit")	Bureau later (specify the general nature of the indications, e.g., "Accession						
For receiving Office use only	For International Bureau use only						
This sheet was received with the international application	This sheet was received by the International Bureau on:						
Authorized officer	Authorized officer						

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

Applicant's or agent's file	25001PCT	International application) Unassigned		
reference number		,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		14.0 (4.0 + 2.1)	***

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred on page 82, line N/A	
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet 🔀
Name of depositary institution American Type Culture Col	lection
Address of depositary institution (including postal code and counting 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	(ng)
Date of deposit April 4, 1997	Accession Number 97976
C. ADDITIONAL INDICATIONS (leave blank if not applicab	This information is continued on an additional sheet
In respect to those designations in which a European Patent is made available until the publication of the mention of the grapplication has been refused or withdrawn or is deemed to be nominated by the person requesting the sample (Rule 28 (4)) D. DESIGNATED STATES FOR WHICH INDICATION	ant of the European patent or until the date of which e withdrawn, only by the issue of such a sample to an expert EPC).
TO STATE THE PURCHASE OF INDICATIONS 4	Mush if not applicable)
E. SEPARATE FURNISHING OF INDICATIONS (leave The indications listed below will be submitted to the International Number of Deposit")	Bureau later (specify the general nature of the indications, e.g., "Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

erri live we we want

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

Applicant's or agent's file	PS001PC7	International application	Una	assigned	
reterence number		0.00	****	27.76	and the first of the second second
		•			

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred on page 76, line N/A							
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet 🔀						
Name of depositary institution American Type Culture Coll	lection						
Address of depositary institution (including postal code and country 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	רצי)						
Date of deposit February 26, 1997	Accession Number 97902						
C. ADDITIONAL INDICATIONS (leave blank if not applicable	le) This information is continued on an additional sheet						
made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC). D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)							
E CEDADATE FUDNICIUMO OF INDICATIONS A	II. billion and Early						
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable) The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")							
For receiving Office use only	For International Bureau use only						
This sheet was received with the international application	This sheet was received by the International Bureau on:						
Authorized officer	Authorized officer						

L. C. Granish

		í
		•

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6:								
16/18,	C12Q	1/68,	G011	C07K 14/47, N 33/50, 33/53,				
33/00,	A61K	20/1/	_					

(11) International Publication Number: **A3**

WO 98/39446

(43) International Publication Date: 11 September 1998 (11.09.98)

(21) International Application Number: PCT/US98/04482

(22) International Filing Date: 6 March 1998 (06.03.98)

(30) Priority Data:

HOLITY DEER.		
60/040,162	7 March 1997 (07.03.97)	US
60/040,333	7 March 1997 (07.03.97)	US
60/038,621	7 March 1997 (07.03.97)	US
60/040,161	7 March 1997 (07.03.97)	US
60/040,626	7 March 1997 (07.03.97)	US
60/040,334	7 March 1997 (07.03.97)	US
60/040,336	7 March 1997 (07.03.97)	US
60/040,163	7 March 1997 (07.03.97)	US
60/043,580	11 April 1997 (11.04.97)	US
60/043,568	11 April 1997 (11.04.97)	US

(Continued on the following page)

(71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). FISCHER, Carrie, L. [US/US]; 5810 Hall Street, Burke, VA 22015 (US). SOP-PET, Daniel, R. [US/US]; 15050 Stillfield Place, Centreville, VA 22020 (US). CARTER, Kenneth, C. [US/US]; 11601 Brandy Hall Lane, North Potomac, MD 20878 (US). BED-NARIK, Daniel, P. [US/US]; 8822 Blue Sea Drive, Columbia, MD 21046 (US). ENDRESS, Gregory, A. [US/US]; 9729 Clagett Farm Drive, Potomac, MD 20854 (US). YU, Guo-Liang [CN/US]; 13524 Straw Bale Lane, Darnestown, MD 20878 (US). NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). FENG, Ping [CN/US]; 4 Relda Court. Gaithersburg, MD 20878 (US). YOUNG, Paul, E. [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). GREENE, John, M. [US/US]: 872 Diamond Drive. Gaithersburg, MD 20878 (US). FERRIE, Ann., M. [US/US]; 13203 L Astoria Hill Court, Germantown, MD 20874 (US). DUAN, Roxanne [US/US]; 4541 Fairfield Drive, Bethesda,

MD 20814 (US). HU, Jing-Shan [CN/US]; 1247 Lakeside Drive #3034, Sunnvvale, CA 94086 (US), GRAVES, Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place #24, Gaithersburg, MD 20878 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne Terrace #316, Gaithersburg, MD 20878 (US). BREWER, Laurie, A. [US/US]; 14920 Mt. Nebo Road, Poolesville, MD 20837 (US). MOORE, Paul, A. [GB/US]; Apartment #104, 1908 Holly Ridge Drive, McLean, VA 22102 (US). SHI, Yanggu [CN/US]; 437 West Side Drive, Gaithersburg, MD 20878 (US). LAFLEUR, David, W. [US/US]; 1615 Q Street, N.W. #807, Washington, DC 20009 (US). LI, Yi [CN/US]; 1247 Lakeside Drive #3034, Sunnyvale, CA 94086 (US). ZENG, Zhizhen [CN/US]; 13950 Saddleview Drive, Gaithersburg, MD 20878 (US). KYAW, Hla [BU/US]: 520 Sugarbush Circle, Frederick, MD 21703 (US).

(74) Agents: BROOKES, Anders, A. et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 10850 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

With an indication in relation to a deposited microorganism furnished under Rule 13bis separately from the description.

Date of receipt by the International Bureau:

06 April 1998 (06.04.98)

(88) Date of publication of the international search report:

23 December 1998 (23.12.98)

(54) Title: 70 HUMAN SECRETED PROTEINS

(57) Abstract

1 - 4

The present invention relates to 70 novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.

60/043,314	11 April 1997 (11.04.97)	US	60/047,598	23 May 1997 (23.05.97)	US	60/056,882	22 August 1997 (22.08.97)	US
60/043,569	11 April 1997 (11.04.97)	US	60/047,613	23 May 1997 (23.05.97:	US	60/056,637	22 August 1997 (22.08.97)	US
60/043,311	11 April 1997 (11.04.97)	US	60/047,582	23 May 1997 (23.05.97)	US	60/056,903	22 August 1997 (22.08.97)	US
60/043,671	11 April 1997 (11.04.97)	US .	60/047,596	23 May 1997 (23.05.97)	US	60/056,888	22 August 1997 (22.08.97)	US
60/043,674	11 April 1997 (11.04.97)	US	60/047,612	23 May 1997 (23.05.97)	US	60/056,879	22 August 1997 (22.08.97)	US
60/043,669	11 April 1997 (11.04.97)	US	60/047,632	23 May 1997 (23.05.97)	US	60/056,880	22 August 1997 (22.08.97)	US
60/043,312	11 April 1997 (11.04.97)	US	60/047,601	23 May 1997 (23.05 97)	US	60/056,894	22 August 1997 (22.08.97)	US
60/043,313	11 April 1997 (11.04.97)	US	60/047,595	23 May 1997 (23.05 97)	US	60/056,911	22 August 1997 (22.08.97)	US
60/043,672	11 April 1997 (11.04.97)	US .	60/047,599	23 May 1997 (23.05 97)	US	60/056,636	22 August 1997 (22.08 97)	US
60/043,315	11 April 1997 (11.04.97)	US	60/047,588	23 May 1997 (23.05 97)	US	60/056,874	22 August 1997 (22.08 97)	US
60/043,578	11 April 1997 (11.04.97)	US	60/047,585	23 May 1997 (23.05 97)	US	60/056,910	22 August 1997 (22.08 97)	US
60/043,576	11 April 1997 (11.04.97)	US	60/047,586	23 May 1997 (23.05.97)	US	60/056,864	22 August 1997 (22.08 97)	US
60/043,670	11 April 1997 (11.04.97)	US	60/047,590	23 May 1997 (23.05 97)	US	60/056,631	22 August 1997 (22.08.97)	US
60/047,600	23 May 1997 (23.05.97)	US	60/047,594	23 May 1997 (23.05 97)	US	60/056,845	22 August 1997 (22 08 97)	US
60/047,615	23 May 1997 (23.05.97)	US	60/047,589	23 May 1997 (23.05 97)	US	60/056,892	22 August 1997 (22 08.97)	US
60/047,597	23 May 1997 (23.05.97)	US	60/047,593	23 May 1997 (23.05 97)	US	60/056,632	22 August 1997 (22 08 97)	US
60/047,502	23 May 1997 (23.05.97)	US	60/047,614	23 May 1997 (23.05.97)	US	60/056,664	22 August 1997 (22.08.97)	US
60/047,633	23 May 1997 (23.05.97)	US	60/047,501	23 May 1997 (23.05.97)	US	60/056,876	22 August 1997 (22 08 97)	US
60/047,583	23 May 1997 (23.05.97)	US	60/048,974	06 June 1997 (06.06.97)	US	60/056,881	22 August 1997 (22 08.97)	US
60/047,617	23 May 1997 (23.05.97)	US	60/048,964	06 June 1997 (06.06.97)	US	60/056,909	22 August 1997 (22 08 97)	US
60/047,618	23 May 1997 (23.05.97)	US	60/056,886	22 August 1997 (22.08.97)	US	60/056,875	22 August 1997 (22 08 97)	US
60/047,503	23 May 1997 (23.05 97)	US	60/056,877	22 August 1997 (22.08.97)	US	60/056,862	22 August 1997 (22 08 97)	US
60/047,592	23 May 1997 (23.05.97)	US	60/056,889	22 August 1997 (22.08.97)	US	60/056,887	22 August 1997 (22 08 97)	US
60/047,581	23 May 1997 (23.05.97)	US	60/056,893	22 August 1997 (22.08.97)	US	60/056,908	22 August 1997 (22 08.97)	US
60/047,584	23 May 1997 (23.05.97)	US		22 August 1997 (22.08.97)	US	60/056,884	22 August 1997 (22.08.97)	US
	23 May 1997 (23.05.97)	US	60/056,878	22 August 1997 (22.08.97)	US	60/057,761	05 September 1997 (05.09.97)	US
	23 May 1997 (23.05.97)	US		22 August 1997 (22.08.97)	US	60/057,650	05 September 1997 (05.09.97)	US
60/047.492	23 May 1997 (23.05.97)	US	60/056.872	22 August 1997 (22.08.97)	US			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spair	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
ΑT	Austria	FR	France	LU	Luxembourg	SN	Senega'
ΑU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monacc	TD	Chac
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistar
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistar.
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Ma!	TT	Trinidad and Tobago
ВJ	Benin	1E	Ire land	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israe!	MR	Mauritania	UG	Uganda
BY	Beiarus	18	Iceland	MW	Malaw:	US	United States of Americ
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistar.
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbahwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	rc	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudar		
DK	Denmark	LK	Sri Lanka	SE	Sweder		
EE	Estonia	LR	Liberia	SG	Singapore		

Asburan N. Contags

PCT/US 98/04482

A CLASSIFICATION OF SUBJECT MATTER IPC 6 C12N15/12 C12 C07K16/18 C12N5/10 C12N1/21 C07K14/47 A61K38/17 G01N33/68 G01N33/50 G01N33/53 C12Q1/68 According to International Patent Classification (IPC) or to both national classification and IPC B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) C12N C07K C12Q G01N A61K Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practical, search terms used) C. DOCUMENTS CONSIDERED TO BE RELEVANT Relevant to claim No. Citation of document, with indication, where appropriate, of the relevant passages Category ° L. HILLIER ET AL.: "The WashU-Merck EST 1-3, X 7-10,21 Project 1997" EMBL SEQUENCE DATABASE, 6 March 1997, HEIDELBERG, FRG, XP002068123 zr78g10.rl Soares NhHMPu S1 Homo sapiens cDNA clone 669570 5' similar to SW: FUCO RAT P17164 Alpha-L-fucosidase precursor; Accession. Accession no. AA234924; -/--Х Patent family members are listed in annex Further documents are listed in the continuation of box C. X Special categories of cited documents : "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to filing date involve an inventive step when the document is taken alone "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another "Y" document of particular relevance; the claimed invention citation or other special reason (as specified) cannot be considered to involve an inventive step when the document is combined with one or more other such docu-ments, such combination being obvious to a person skilled in the art. "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but "&" document member of the same patent family later than the priority date claimed Date of mailing of the international search report Date of the actual completion of the international search 1 6. 09. 1998 16 June 1998 Authorized officer Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentiaan 2 NL - 2280 HV Rijawijk Tel. (+31-70) 340-2040, Tx. 31 651 epo ni, HORNIG H. Fax: (+31-70) 340-3016

Internal Application No PC1/US 98/04482

		PC1/US 98/04482		
	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	Relevant to claim No.		
Category °	Citation of document, with indication, where appropriate, of the relevant passages	riemyani le ciami iyo.		
X	L. HILLIER ET AL.: "The WashU-Merck EST Project" EMBL SEQUENCE DATABASE, 15 December 1996, HEIDELBERG, FRG, XP002068124 z140b11.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 504381 5' similar to TR:G182779 Lysosomal Enzyme Alpha-L-Fucosidase Accession no. AA151194	1-3, 7-10,21		
X	L. HILLIER ET AL.: "The WashU-Merck EST Project" EMBL SEQUENCE DATABASE, 4 June 1996, HEIDELBERG, FRG, XP002068125 zc54a02.rl Soares senescent fibroblasts NbHSF Homo sapiens cDNA clone 326090 5' similar to SW:FUCO_HUMAN P4066 tissue Alpha-L-Fucosidase precursor; Accession no. W52490	1-3, 7-10,21		
A	WO 97 07198 A (GENETICS INSTITUT) 27 February 1997 see the whole document	1-23		
A	WO 97 04097 A (GENETICS INST) 6 February 1997 see the whole document	1-23		
A	US 5 536 637 A (JACOBS KENNETH) 16 July 1996 see the whole document	1-23		
Α .	JACOBS K ET AL: "A novel method for isolating eukaryotic cDNA clones encoding secreted proteins." KEYSTONE SYMPOSIUM ON DENDRITIC CELLS: ANTIGEN PRESENTING CELLS OF T AND B LYMPHOCYTES, TAOS, NEW MEXICO, USA, MARCH 10-16, 1995. JOURNAL OF CELLULAR BIOCHEMISTRY SUPPLEMENT 0 (21A). 1995. 19. ISSN: 0733-1959, XP002027246 abstract no. C1-207 see abstract	1-23		
A	WO 90 14432 A (GENETICS INST) 29 November 1990 see the whole document	1-23		
Α	WO 96 17925 A (IMMUNEX CORP) 13 June 1996 see the whole document	1-23		

Antonio (1994) Original Andreas

Inter onal Application No PC I / US 98/04482

		PC1/US 98/04482		
C.(Continua	.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT			
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.		
A	T. OCCHIODORO ET AL.: "Human alpha-L-Fucosidase: Complete coding sequence from cDNA clones" BIOCHEM. AND BIOPHYS. RES. COMMUNICATIONS, vol. 164, no. 1, 16 October 1989, ACADEMIC PRESS, NEW YORK, US, pages 439-445, XP002068126 cited in the application see the whole document	1-23		
ļ				

And Committee of the Co

rnational application No.

INTERNATIONAL SEARCH REPORT

PCT/US 98/04482

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)	
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:	
1. X Claims Nos because they relate to subject matter not required to be searched by this Authority, namely Remark: Although claim 17 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.	
Claims Nos. because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:	
3. Claims Nos because they are dependent claims and are not crafted in accordance with the second and third sentences of Rule 6.4(a).	
Box II Observations where unity of Invention is lacking (Continuation of Item 2 of first sheet)	
This International Searching Authority found multiple inventions in this international application, as follows:	
see further information sheet	
As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims	
As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.	
3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:	
A. X No required additional search fees were timely paid by the applicant. Consequently, this international Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos. See further information sheet	
Remark on Protest The additional search fees were accompanied by the applicant's protest No protest accompanied the payment of additional search fees.	

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1. Claims: (1-23) partially

-An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence consisting of SEQ ID no. 11; wherein said polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein encoding the sequence of SEQ ID no. 134 or the polypeptide encoded by the cDNA sequence included in ATCC Deposit no: HGCMD20, which is hybridizable to SEQ ID no.11; a recombinant vector comprising said isolated nucleic acid molecule; a method of making a recombinant host cell comprising said isolated nucleic acid molecule; a recombinant host cell comprising said vector; an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence consisting of SEQ ID no. 134; an isolated antibody that binds specifically to said isolated polyeptitde; a recombinant host cell that expresses said isolated polypeptide; a method of making said polypeptide; a method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of said polypeptide; a method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject using said polynucleotide and/or polypeptide sequences; a method for identifying a binding partner to said polypeptide; a gene corresponding to the cDNA sequences of SEQ ID no.11; a method for identifying an activity in a biological assay, by using the expression of SEQ ID no. 134;

Inventions 2 to 70. Claims: (1-23) partially

-Idem as subject 1 but limited to gene nos. 2 to 70 respectively cDNA clone sequences HLDBG33 to HMCAB89. (Invention 2 is limited to SEQ ID nos.12,81,135, and 204; Invention 3 is limited to SEQ ID nos.13 and 136;; Invention 70 is limited to SEQ ID nos.80 and 203;)

mation on patent family members

Internal Application No
PC7 / US 98/04482

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9707198 A	27-02-97	US 5707829 A AU 6712396 A AU 6768596 A EP 0839196 A EP 0851875 A WO 9704097 A	13-01-98 18-02-97 12-03-97 06-05-98 08-07-98 06-02-97
WO 9704097 A	06-02-97	US 5707829 A AU 6712396 A EP 0839196 A AU 6768596 A EP 0851875 A WO 9707198 A	13-01-98 18-02-97 06-05-98 12-03-97 08-07-98 27-02-97
US 5536637 A	16-07-96	US 5712116 A	27-01-98
WO 9014432 A	29-11-90	US 5580753 A AT 147436 T AU 637620 B AU 5928990 A CA 2056997 A DE 69029657 D DK 473724 T EP 0473724 A ES 2099096 T JP 4506006 T US 5734037 A US 5414071 A	03-12-96 15-01-97 03-06-93 18-12-90 24-11-90 20-02-97 14-04-97 11-03-92 16-05-97 22-10-92 31-03-98 09-05-95
WO 9617925 A	13-06-96	AU 4639396 A CA 2206488 A FI 972390 A NO 972455 A	26-06-96 13-06-96 05-06-97 06-08-97